CHARACTERIZATION OF THE IN VITRO TRANSCRIPTION OF CHROMATIN

by

Ruth Anne Gjerset
B.A., University of California, San Diego, 1971

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOCHEMISTRY

in the

GRADUATE DIVISION

(San Francisco)

of the

UNIVERSITY OF CALIFORNIA

Approved: Brian J. McCarthy
Christine Guthrie
James A. Spudich
Committee in Charge

Deposited in the Library, San Francisco Medical Center:

.....
Date Librarian

JAN 2 1977

Degree Conferred: .....

ABSTRACT

Chromatin from Schneider's *Drosophila melanogaster* tissue culture cells and from mouse L-cells was transcribed *in vitro* using RNA polymerase purified from *E. coli*. The experiments were designed to characterize the overall specificity of *in vitro* transcription, as well as the fidelity with which the enzyme selects against regions which are not transcribed *in vivo*. These non-transcribed regions include anti-strand and satellite sequences. The overall specificity of transcription was monitored from Drosophila chromatin using a probe representing a substantial portion of the sequences transcribed *in vivo*, namely a cDNA copy of poly(A)-containing nuclear RNA. This probe was also used to monitor the frequency distribution of sequences within the *in vitro* transcript. The concentration of anti-strand sequences in the Drosophila chromatin transcript was monitored using labeled nuclear RNA as a probe. The fidelity with which the *E. coli* enzyme selects against satellite sequences was determined using mouse L-cell chromatin prepared by several procedures and transcribed under several sets of conditions. The transcripts were then assayed for the presence of satellite sequences using labeled satellite DNA as a probe. Taken together, the results showed that *E. coli* RNA polymerase preferentially transcribes from regions of chromatin which are transcribed *in vivo*. In addition, sequences which are abundant *in vivo* are transcribed more frequently *in vitro*. However, the enzyme also transcribes some anti-strand and satellite sequences. Changes in the methods for chromatin isolation or transcription failed to eliminate this aberrant transcription. However, when mouse chromatin was transcribed at lower

enzyme to template ratios, satellite sequence transcription represented a smaller proportion of the total transcription, suggesting that these regions possess weaker initiation sites for the enzyme than do those regions which are transcribed in vivo. Aberrant transcription is therefore due at least in part to the high enzyme to template ratios used during transcription but may also be due to the loss of some structural integrity of chromatin during isolation or to the inability of the bacterial enzyme to properly recognize eucaryotic promoters.

TO MY PARENTS

ACKNOWLEDGEMENTS

I take this opportunity to say thanks to each person who has offered his or help and support to me during my career here as a graduate student. I feel fortunate to have worked in an environment which offers such excellent science. To my advisor, Brian McCarthy, I express my sincere appreciation for his guidance, help and patience throughout my graduate work. I am also indebted to Harald Biessmann with whom I collaborated during much of this work. In addition to them, several other people have helped in the preparation of this thesis, and I am happy to be able to thank them for the interest they took and for the time and care they spent in critically reading it. These people are Christine Guthrie and Jim Spudich, members of my thesis committee, and Walter Miller, my friend and colleague. And thanks also to Leslie Spector, who was a tremendous help in preparing the final copy. To Beatriz Levy W., who shared many of the joys and traumas of graduate school with me, go my special thanks for her constant support, encouragement and friendship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

DNA	deoxyribonucleic acid
RNA	ribonucleic acid
mRNA	messenger RNA
rRNA	ribosomal RNA
HnRNA	heterogeneous nuclear RNA
cDNA	complementary DNA
^{3}H	tritium
ATP	adenosine triphosphate
GTP	guanosine triphosphate
CTP	cytidine triphosphate
UTP	uridine triphosphate
Hg-UTP	5 mercuriuridine triphosphate
poly(A)	polyadenosine
$R_o t$	product of initial RNA concentration (moles of nucleotide per liter) and time (seconds)
$C_o t$	product of initial DNA concentration (moles per liter) and time (seconds)
l	liter
ml	milliliter
M	molar (moles per liter)
mM	millimolar
μM	micromolar
g	gram
mg	milligram

µg	microgram
O.D.	optical density
nm	nanometer
NP-40	Nonidet P-40
TCA	trichloroacetic acid, $CHCl_3COOH$
SDS	sodium dodecyl sulfate
EDTA	ethylene diaminetetraacetic acid
DTT	dithiothreitol
SSC	0.15 M NaCl, 0.015 M Na-citrate
RSB	reticulocyte standard buffer
PEB	phosphate equimolar buffer
K_m	Michaelis-Menten constant
TNE	Tris, NaCl, EDTA buffer

A. INTRODUCTION

One problem basic to the molecular biology of eucaryotes concerns the mechanism by which cells derived from a common precursor cell become determined to different developmental pathways and how they achieve the controlled and selective expression of their genetic material. From various lines of evidence it appears that the DNA in all somatic cells of an organism is the same both in quantity and in the kinds of sequences present. That the quantity of DNA per nucleus is the same has been shown by Mursky and Ris (1951) and Vendreley (1955). That the same sequences are present in all cell types is inferred from nuclear transplantation experiments in Xenopus where the genetic information for an entire organism could be derived from a single somatic cell (Gurdon, 1962), and from DNA annealing experiments which failed to detect base sequence differences between various tissues (McCarthy and Hoyer, 1964). Therefore, differences in genetic expression amongst different somatic cells of an organism do not appear to be a manifestation of different DNA contents and this implies the existence of mechanisms to regulate which genes are expressed.

The details of these mechanisms are not understood, however. It is not possible to make straightforward analogies between procaryotic and eucaryotic systems because the genome of eucaryotes is distinguished from procaryotes in several fundamental respects. First, it is up to one thousand times larger (as in the case of mice compared to E. coli) and therefore several orders of magnitude larger than it need be to code for all of the known enzymes. Because of this, one might suppose that only a small fraction of genome is transcribed. Indeed, it

appears from nucleic acid hybridization that this is the case, in contrast to the situation in procaryotes where nearly all (80-90% of one strand) of the DNA is transcribed (McCarthy and Bolton, 1964; Kennell, 1968). A second distinguishing feature of the eucaryotic genome is the presence of reiterated or repetitive sequences. These are sequences which are present in multiple copies, or are similar enough to renature with each other. Their function is not known, however, and their amount per genome varies markedly between various organisms (15% in mice, 15% in Drosophila, 80% in Amphiuma, 45% in Xenopus). Because of this, there may be no correlation, as seen in procaryotes, between genome size (the total amount of DNA per haploid genome equivalent) and genetic complexity (the amount of genomic DNA in diverse, non homologous, i.e., unique sequences). In plant species classified as Ranunculaceae, for example, there exists an 80-fold variation in DNA content among various species of quite similar phenotype.

Early attempts to characterize the nature of genetic regulation in eucaryotes were aimed at quantitating and comparing the fraction of genes transcribed in various cells and tissues of an organism. An excess of cellular RNA was hybridized in increasing amounts with purified unique sequences of DNA. The fraction of DNA hybridized at saturation was an indication of the fraction of the genome transcribed. Such data support the notion that most of the eucaryotic genome is not transcribed at any given time. For example, 1.2% of the unique sequence DNA can be hybridized with Xenopus oocyte RNA (Davidson and Hough, 1971), 12% in newborn mice (Gelderman et al., 1971), 10% in mouse brain (Hahn and Laird, 1971; Brown and Church, 1972), and 4-5% in

mouse kidney, liver or spleen (Grouse et al., 1972). Although there is partial overlap of sequences, the populations of RNA differ from tissue to tissue in mammals (Grouse et al., 1972) and from one developmental stage to another for Dictyostelium (Firtel, 1972) and Drosophila (Turner and Laird, 1973). It is important to note however, that final saturation levels cannot be determined precisely, since saturation is approached asymtotically as RNA concentration increases. Since there are theoretical limits to the maximum RNA concentration attainable, it is possible that this approach is not sensitive to very rare RNA species, which would hybridize only at high RNA concentrations or after very long annealing times. Therefore, these are minimum estimates of the amount of the genome transcribed.

A more recent approach to this question attempts to avoid the limitations of unique sequence DNA hybridizations to excess RNA. This is accomplished by using a poly(A)-containing population of cellular RNA as a template for a complementary DNA (cDNA) copy synthesized by reverse transcriptase. Hybridizations are then performed between the cDNA and an excess of its template. Since, theoretically, all of the cDNA can be hybridized, these reactions offer an advantage over those employing unique sequence DNA by giving a clear indication as to when the reaction reaches completion. The kinetics of the reaction allow one to determine the number of different sequences present in the RNA population (that is, the complexity of the population). Using this approach, the complexity of nuclear and messenger RNA populations has been analyzed in several systems including HeLa cells (Bishop, 1974), mouse, brain, kidney, and L-cells (Ryffel and McCarthy, 1975), mouse Friend cells (Getz et al., 1975), Schneider's Drosophila

cells (Levy W. and McCarthy, 1975; Levy W. et al., 1976), and chick oviduct and liver (Axel et al., 1976). In all cases it appears that only a fraction of the DNA sequences are represented in nuclear or cytoplasmic poly(A)-containing RNA. In addition, this method has revealed in all cases examined, that different sequences are represented at different frequencies within the same population. This conclusion implies the presence of quantitative as well as qualitative cellular controls.

The diversity of RNA sequences from one cell type to another in an organism therefore contrasts with the invariability of DNA sequences and suggests that a fundamental mode of gene regulation is transcriptional. The notion that the proteins of chromatin are responsible for transcriptional regulation in eucaryotes has provided the impetus for developing in vitro systems for studying the possible regulatory role of the chromosomal components and for investigating the molecular interactions responsible for transcriptional control. Early studies of in vitro transcription showed that chromatin is a poor substrate for E. coli RNA polymerase compared to DNA (Bonner et al., 1968; Cedar and Felsenfeld, 1973) and that this inefficiency is due largely to a decrease in the number of initiation sites for RNA polymerase on chromatin (Ceder and Felsenfeld, 1973).

Initial attempts to determine the specificity of in vitro transcription from chromatin used filter hybridization techniques involving competition of various RNA populations for hybridization to immobilized DNA or hybridization of a sample RNA population to DNA filters presaturated with a reference RNA population. These techniques are subject

to a strict limitation in interpretation because the hybridization conditions (that is, the amount of DNA which could be fixed to the filter) allowed the detection only of RNA sequences which hybridize with repetitive DNA. Nevertheless, they were a useful means for monitoring the specificity of the in vitro product and the results of these studies suggested that even with E. coli RNA polymerase, a fair degree of transcriptional specificity occurred in vitro and that this specificity paralleled the in vivo specificity. In several tissues of rat, for example (Paul and Gilmour, 1968; Tan and Miyagi, 1970) and rabbit (Paul and Gilmour, 1968) template restriction was observed in vitro using E. coli RNA polymerase (3-10% of the transcription from naked DNA of the same source) and the restriction was tissue specific. No differences could be detected by this method between in vivo and in vitro populations of RNA. Smith et al. (1969) showed tissue specific restriction with mouse kidney, liver and spleen using either E. coli or mouse polymerases, and failed to detect differences between in vivo and in vitro populations. Even after salt and urea dissociation and reconstitution, tissue specific transcription seemed to occur in vitro (Bekhor et al., 1969; Huang and Huang, 1969).

A more sensitive approach seeks to characterize the transcript by measuring the concentration of a specific RNA base sequence in the in vitro transcript. In this approach one monitors the hybridization with a labeled probe (cDNA) complementary to the sequence of interest. Several systems have been examined in this manner including those for globin (Paul et al., 1973; Axel et al., 1973; Steggles et al., 1974) in which it was found that globin-specific sequences were present in transcripts from erythropoietic chromatin, but were at least an order

of magnitude lower in concentration in chromatin transcripts from other tissues and from naked DNA using E. coli RNA polymerase or eucaryotic RNA polymerase; the SV3T3 system (Astrin, 1973, 1975; Shih et al., 1975) where the same portion of the early strand was transcribed both in vivo and in vitro; the histone system (Stein et al., 1975) in which transcripts of HeLa S phase but not G1 phase cells contained histone specific sequences; the murine leukemia virus system (Janowski et al., 1974) in which it was shown that although normal and leukemic mice both contained viral sequences in their spleen DNA, viral transcription was enhanced in chromatin isolated from leukemic mouse spleens as compared to chromatin from normal mouse spleens; the immunoglobulin system (Smith and Huang, 1976) in which kappa chain message transcription both from nuclei with endogenous polymerase and chromatin with E. coli polymerase was found to be more efficient with nuclei or chromatin from kappa chain producing cells than from lambda chain producing cells; and the ovalbumin system (Harris et al., 1976) in which ovalbumin sequences were detectable in transcripts of chromatin from estrogen stimulated oviducts, but not from unstimulated or hormone-withdrawn oviduct chromatin or spleen chromatin and to a much lower degree (10% of level in stimulated chromatin) from naked DNA. Marzluff and Huang (1975) have also reported that 5S and tRNA species could be transcribed in vitro from mouse myeloma chromatin using endogenous RNA polymerase with correct strand selection and with reinitiation. Therefore it seems from these studies that the presence or absence of a particular sequence in both the E. coli and mammalian enzyme transcripts correlates well with the source of the chromatin.

Contrasting with these results are the results of other experiments designed to assay for the transcription of sequences known to be repressed *in vivo*. These include studies of ribosomal and 5S RNA in *Xenopus laevis* chromatin (Reeder, 1973; Honjo and Reeder, 1974) in which both strands as well as spacer regions were transcribed whether *E. coli* or *Xenopus* polymerases I and II were used. In mouse chromatin, Reeder (1973) detected significant levels of satellite DNA transcription (about one-tenth the level from naked DNA) using filter hybridizations. Similarly, *E. coli* RNA polymerase or sheep RNA polymerase transcribed sequences from the anti strand of the globin gene in rabbit bone marrow-chromatin (Wilson *et al.*, 1975).

Therefore, despite numerous studies aimed at characterizing *in vitro* transcription of chromatin, there remains some doubt as to the overall fidelity of the process. Conclusions drawn from filter hybridization experiments apply only to the transcription from repetitive DNA sequences. The experiments with specific probes suffer from the limitation that they leave the greater portion of the transcript uncharacterized. The experiments which examine aberrant transcription leave open the possibility that variations in the method of chromatin preparation or variations in the conditions of transcription would influence the fidelity with which *E. coli* RNA polymerase transcribes. The present work was undertaken in an attempt to resolve some of these questions. Using chromatin from *Drosophila melanogaster* tissue culture cells and *E. coli* RNA polymerase, I have asked to what extent the overall specificity of transcription is retained *in vitro* with respect to 1) the kinds of sequences which are transcribed, and 2) the frequency with which these sequences are represented in the transcript. For this, a general probe

was used, namely a cDNA copy of poly(A)-containing nuclear RNA. This probe represents a substantial portion (80-90%) of the sequence types present *in vivo*. I have also asked to what extent various chromatin isolation procedures and transcription conditions affect the level of aberrant transcription. For these experiments, L-cell chromatin was transcribed and the transcript was assayed for the presence of satellite sequences, which are not present in *in vivo* RNA.

The *Drosophila* system offers numerous advantages for biochemical studies. It is one of the most extensively studied eucaryotic systems and offers the investigator the chance to draw on the studies from diverse disciplines, including genetics (Lindsley and Grell, 1972) cytology and cytogenetics (Pardue *et al.*, 1970). One of the greatest advantages for hybridization studies of the sort attempted here is the size of its genome, about 10^{11} daltons (Rasch *et al.*, 1971; Laird, 1971) making it the smallest known metazoan genome (Shapiro, 1970) of which 75% is unique (McConaughy *et al.*, 1969; Wu *et al.*, 1972). Therefore the hybridization reactions can be completed at much lower $C_o t$ values. It has also been the subject of studies on the organization of the eucaryotic chromosome (Peacock *et al.*, 1973; Laird *et al.*, 1973; Rudkin and Tartof, 1973), studies which will undoubtedly lead to insights into gene regulation and serve as a complement to biochemical studies on RNA populations.

A further advantage which facilitates biochemical analysis, is the presence of several lines of Drosophila cells (Kakpakov *et al.*, 1969; Echalier *et al.*, 1970; Schneider, 1972; Schneider and Blumenthal, 1976). Manipulation of cells in culture is much easier than with the whole organism and more detailed biochemical questions can be phrased. Schneider's

Drosophila melanogaster line 2 cells were used for these experiments, because of the data available concerning nuclear and cytoplasmic RNA populations and the relationship between the two. Earlier studies showed that, as with larvae, pupae and adult flies, a large fraction of the cellular DNA is transcribed in tissue culture, about 25% of unique sequence DNA (Turner and Laird, 1973; McCarthy *et al.*, 1973). These findings have been extended using cDNA copies of poly(A)-containing nuclear RNA populations (Levy W. *et al.*, 1976), and allow the determination of the number of different sequences in the RNA as well as the relative abundances of these sequences. Such studies confirm earlier findings regarding the fraction of the genome transcribed and show that the majority of this transcription (more than 80%) is accounted for by nuclear poly(A)-containing species. This nuclear population is a complex one in that some sequences in the population occur much more frequently than do others. We therefore have a highly characterized probe for the overall specificity of *in vitro* transcription from chromatin isolated from these cells, namely the cDNA complementary to poly(A)-containing nuclear RNA. Because the nuclear RNA population contains a broad frequency distribution of sequences, it was possible to fractionate the cDNA by partial annealing to nuclear RNA so as to obtain probes enriched for frequent or rare nuclear RNA sequences. These probes provide an additional tool for monitoring *in vitro* transcription allowing comparison of the frequency distribution *in vivo* and *in vitro* RNA populations.

Chromatin from mouse L-cells was used to examine aberrant transcription. These studies are intended to serve as a complement to the studies with Drosophila cell chromatin in that in the latter I probe for

sequences which should be transcribed and here I probe for sequences which should be repressed. Mouse satellite DNA comprises about 10% of the genome and reassociates so rapidly following thermal denaturation that Waring and Britten (1966) estimated that it must consist of about 10^6 repeats of a sequence about 350 bp long. Flamm et al. (1969) have shown that satellite sequences are repressed in vivo to less than one part in 60,000 of the total cellular RNA. The unusual base composition of these sequences relative to bulk DNA and the slight enrichment of purines in one strand makes it possible to isolate satellite DNA on density gradients and separate the two DNA strands (Lieberman, 1973; Flamm et al., 1969). Mouse satellite DNA therefore provides a convenient probe for transcriptionally inactive chromatin.

B. MATERIALS AND METHODS

1. Cells

Schneider's line 2 cells, a line derived from early embryos of Drosophila melanogaster (Schneider, 1972) were grown at 25°C in roller bottles, plastic Falcon flasks, or suspension culture, using Schneider's tissue culture media (Pacific Biologicals, Co.) supplemented with 5 mg/ml bactopeptone, 15% heat inactivated fetal calf serum, streptomycin (100 µg/ml), penicillin (100 units/ml), and nystatin (50 units/ml). Cells growing in roller bottles and Falcon flasks were transferred every 2-3 days by shaking the cells from the surface and diluting with fresh medium. Cells growing in suspension were transferred every 4 days. Cells grew logarithmically from 10^5 cells/ml to 10^7 cells/ml, doubling once a day. For experiments, cells from mid to late log phase growth were used.

L-cells (line A9) were grown in suspension at 37°C in Joklik's modified essential medium containing 5% calf serum and 100 units/ml each of penicillin and streptomycin. Cells were diluted with fresh medium every 2 days and grow logarithmically from 2-8 x 10^5 cells/ml doubling once a day. For experiments, cells from mid to late log phase growth were used.

2. Enzymes

a. DNA dependent RNA polymerase was prepared from E. coli K12 grown in high peptone medium in a fermentator to late log phase growth. The purification procedure was that of Burgess (1969) except that the final glycerol gradient step was omitted. For a typical preparation, about

250 g of cells were obtained from 9 liters of high density culture, and lysed by shearing in a French pressure cell. The cell extract was treated with DNase I and ammonium sulfate as described and applied to a DEAE-cellulose column. The fractions containing enzyme activity were pooled and precipitated with ammonium sulfate. After resuspension and dialysis the enzyme solution (24 ml) was layered onto 6 SW27 gradients, and centrifuged 24 hr at 25K at 4°C. 1.5ml fractions were collected and assayed for O.D. 280 and activity. The fractions containing enzyme activity were pooled, precipitated with ammonium sulfate and resuspended in storage buffer at a concentration of 5 mg/ml as determined by O.D. 280, assuming an extinction coefficient of $E_{280}^{1\%} = 6.5$. The enzyme was stored at -20°C. Enzyme activity assayed on high molecular weight calf thymus DNA, was about 300 units/mg. This is approximately half of the activity reported by Burgess (1969). The enzyme was about 50% pure, as judged by SDS gel electrophoresis. Since no nuclease or protease activity was detected in the enzyme preparation, it was assumed that this contaminant did not influence the transcription results.

b. S1 nuclease was purchased from P.L. Biochemicals or prepared by the method of Sutton (1971).

c. AMV reverse transcriptase was provided by Dr W.J. Rutter and Dr. J. Beard (Life Sciences, Inc., St. Petersburg, Fla) through the courtesy of Dr. M.A. Chirigos of NIH.

d. Pancreatic RNase (Sigma) was resuspended in H_2O at 2 mg/ml and boiled 10 min to destroy residual DNase activity, and stored at -20°C.

e. DNase I (Sigma, from bovine pancreas, RNase free) was resuspended at 1 mg/ml in 10 mM Tris, pH 7.5, 5 mM $MgCl_2$ and stored at -20°C in 0.5 ml aliquots.

3. Chromatin Preparation

For a typical preparation, 1 liter of mid to late log phase Drosophila cells were collected by centrifugation at 1500xg for 10 min. All further manipulations were done at 0-4°C. Cells were then washed once with phosphate buffered saline (PBS) (0.14 M NaCl, 3 mM KCl, 9.5 mM Na_2HPO_4, 1.5 mM KH_2PO_4) and disrupted by homogenization with a tight fitting pestle in a Dounce homogenizer (20 strokes) in 75 ml 0.3 M sucrose, 2 mM Mg acetate, 3 mM CaCl, 10 mM Tris-HCl, pH 8.0; 0.5 mM dithiothreitol was added and the homogenate layered over a cushion of 10 ml of the 2 M sucrose, 5 mM MgAcetate, 10 mM Tris, pH 8, 0.5 mM dithiothreitol. After carefully stirring the interphase, nuclei were collected by centrifugation in a Beckman SW27 rotor at 15,000 rpm for 45 min at 4°C. The nuclear pellet was washed once in 10 mM Tris-HCl, pH 8.0 and collected at 1000xg for 10 min. The nuclei appeared to be free of cytoplasmic contamination as judged by phase contrast microscopy. Nuclei were taken up in 3-5 ml of 10 mM Tris-HCl, pH 8.0 and disrupted by passing through a French pressure cell at 1000 psi. The supernatant after another centrifugation at 1000xg for 10 min in a Sorvall NB4 rotor was used within 24 hr as "chromatin" in the in vitro transcription experiments. One liter of cells usually yielded 1-2 mg chromatin as judged by O.D. 260.

Chromatin from L-cells was prepared as described above, or prepared by the procedure of Bonner et al. (1968) except that chromatin was not pelleted through sucrose. For a typical preparation, 1 liter of mid to late log phase L-cells ($5x10^8$-10^9 cells) were collected by centrifugation at 1500xg for 10 min. This procedure differs from the former in that nuclei are not exposed to Ca^{++} and Mg^{++} cations which are

known to be necessary for the activation of the endogenous endonuclease present in nuclei (Hewish and Burgoyne, 1973). All further manipulations were done at 0-4°C. Cells were washed several times with saline-EDTA, pH 8 (0.075 M NaCl, 0.024 M EDTA, 10 mM Tris, pH 8) and disrupted by homogenization with a loose fitting pestle in a Dounce homogenizer (15 strokes) in 10 ml saline-EDTA containing 0.5% NP40 (Shell Oil Co.). The total incubation time in NP40 was kept to 4-5 min. Nuclei were spun out and washed several times in saline-EDTA. Nuclei were lysed by an abrupt decrease in osmolarity, that is, by homogenizing in a loose fitting Dounce homogenizer with 10 mM Tris, pH 8 and pelleting chromatin at 8000 rpm in HB4 rotor. For solubilization, chromatin was resuspended in 3 ml 10 mM Tris, pH 8 and passed through a French pressure cell at 3000 psi. It was then spun at 10K for 10 min in an HB4 rotor and the pellet was discarded. The supernatant which usually contained 1-2 mg of chromatin (as judged by O.D. 260) was then used within 24 hr for transcription experiments.

4. Reconstitution of Chromatin

Chromatin was reconstituted following the method of Kleiman and Huang (1972). In a typical preparation, 2 mg of chromatin in about 2 ml 10 mM Tris, pH 8 was dissociated by dialyzing overnight against 100 volumes of 5 M urea, 2 M NaCl, 10 mM Tris, pH 8. Reconstitution was accomplished by stepwise reduction of the NaCl concentration to 0.1 M followed by dialysis into 10 mM Tris, pH 8 as follows: 6 hr dialysis against 100 volumes 0.6 M NaCl, 5 M urea, 10 mM Tris, pH 8; 6 hr dialysis against 100 volumes 0.1 M NaCl 5 M urea, 10 mM Tris, pH 8; overnight dialysis against 1000 volumes 10 mM Tris, pH 8; followed by two additional dialyses of 2 hr each against 500 volumes 10 mM Tris, pH 8.

5. Sizing of the DNA in Chromatin

The size distribution of the DNA in chromatin was determined on a 0.9% agarose gel using 0.09 M Tris, 2.5 mM EDTA, 0.09 M boric acid buffer. After electrophoresis the gel was stained with ethidium bromide, photographed and the negative was scanned in a Beckman spectrophotometer.

6. Preparation of DNA for Transcription

Chromatin in 10 mM Tris, pH 8 was made 0.5% in SDS and 0.1 M in NaCl. It was then extracted once with phenol:chloroform (1:1) equilibrated in 10 mM Tris, pH 8, 0.1 M NaCl and several times with chloroform + 4% isoamyl alcohol. The aqueous phase was then ethanol precipitated and the precipitate was resuspended in 0.1xSSC (1xSSC = 0.15 M NaCl, 0.015 M NaCitrate). Pancreatic RNase was added to a final concentration of 10 µg/ml and the solution was incubated at 37°C for 30 min. 0.5% SDS was added and the nucleic acids were re-extracted and ethanol precipitated. The precipitate was resuspended in 10 mM Tris, pH 8 and dialyzed against 100 volumes of the same buffer for several hours and then 100 volumes of fresh buffer for several more hours. The DNA in the final solution was used for transcription or stored at -20°C.

7. Preparation of Mercurated Uridine Triphosphate (Hg-UTP)

Mercurated UTP was synthesized following the procedure of Dale et al. (1975). A solution of 20 mM UTP in 0.1 M sodium acetate, pH 6 was incubated at 50°C with the same volume of freshly prepared 0.1 M mercuric acetate in the same buffer. Most of the unreacted mercury ions were removed by passing the reaction mixture over a Sephadex G-10 column equilibrated in 0.1 M NaAc, pH 6 and the mercurated UTP was further purified by affinity chromatography on DEAE-cellulose (Whatman DE 52).

The compound was batch-eluted with 0.5 M triethylammoniumhydrocarbonate, pH 8, evaporated to dryness and washed three times with methanol. It was suspended at a concentration of 1-4 mg/ml in 10 mM Tris-HCl, pH 7.4 and stored at -20°C. The compound was then subjected to spectral and chromatographic analysis (detailed in Results section).

8. Transcription

This and all subsequent procedures involving RNA were performed with sterile buffers and with glassware which had been heated at least 2 hr at 180°C to destroy RNase. Drosophila or mouse chromatin and DNA were transcribed *in vitro* using essentially the incubation procedure described by Astrin (1975). The 5 ml reaction mixture contained 150 mM KCl, 40 mM Tris-HCl, pH 8, 0.1 mM dithiothreitol, 5 mM $MgCl_2$, 1 mM $MnCl_2$, 4 mM of each ATP, GTP, CTP (Sigma), 1 mM HgUTP, 20 mM 2-mercaptoethanol and 0.5 mg/ml chromatin or DNA as template and 0.8 mg of *E. coli* RNA polymerase per mg template purified according to Burgess (1969) through the first glycerol gradient. In experiments in which the transcript was to be labeled, the CTP concentration was lowered to 1 mM and 30-100 µCi ^{3}H-CTP (23 Ci/mmole, Schwarz/Mann) were added per ml reaction mixture. After 4 hr incubation at room temperature the solution was made 0.5% in SDS and 10 mM in EDTA and extracted as described below.

L-cell chromatin was also transcribed *in vitro* using the method of Axel *et al.* (1973). A 2.5 ml reaction mixture contained 0.75-1 mg chromatin or DNA, 1 mg *E. coli* RNA polymerase, 10 mM Tris-HCl, pH 8, 1 mM $MnCl_2$, 0.08 mM each of ATP, GTP, CTP and UTP. Synthesis continued at 37°C for 45 min and was terminated by the addition of 0.5% SDS and 10 mM EDTA. RNA was extracted as described below.

9. Purification of In Vitro Synthesized RNA

RNA was extracted as described by Palmiter (1974). One volume of water saturated phenol was added to the transcription mixture (which had been made 0.5% in SDS and 10 mM in EDTA) and the mixture was vortexed. Phases were separated by brief centrifugation (1 min, 15,000xg) and the lower phase was removed. The aqueous phase and the interphase were re-extracted 2-3 times with chloroform + 4% isoamyl alcohol, brought to 0.1 M NaCl (in the case of RNA transcribed under low salt conditions) and precipitated with 2 volumes of ethanol overnight at -20°C.

In the case of transcripts prepared in the absence of Hg-UTP, the precipitated nucleic acids were resuspended in 10 mM Tris, pH 7.4, 0.1 M NaCl, 5 mM $MgCl_2$ and freed of template by treatment with 40 µg/ml DNase I for 1 hr at 37°C. The mixture was re-extracted as before, precipitated wtih ethanol and subjected to a second cycle of DNase I treatment. Such treatment was sufficient to degrade greater than 99% of the DNA to acid solubility. The mixture was again extracted and ethanol precipitated. The precipitated nucleic acids were then freed of residual triphosphates and monophoshpates by filtration over a Sephadex G-50 column equilibrated in sterile H_2O. The RNA sample was resuspended in 200 µl sterile H_2O and passed over a 5 ml G-50 column and 1 ml fractions were collected.

In the case of transcripts prepared in the presence of Hg-UTP, the precipitated nucleic acids were dissolved in 1-2 ml of water and made 1x TNE (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA). Triphosphates were removed from the higher molecular weight material by chromatography on a 50 ml Sephadex G-50 equilibrated in sterile TNE. The material in

the excluded volume was collected and precipitated again with 2 volumes of ethanol. After dissolving the precipitate in 1 ml TNE the in vitro synthesized RNA, containing Hg-UTP, was separated from DNA and endogenous RNA by affinity chromatography on sulfhydryl-Sepharose 6B. This resin was prepared according to Cuatrecasas (1970) and equilibrated in TNE. After incubating for 30 min on a 5 ml column the unbound material was passed over the column again and the resin washed extensively with TNE until no more nucleic acid could be removed. The bound RNA was eluted with TNE containing 0.1 M 2-mercaptoethanol and precipitated with 2 volumes of ethanol at -20°C. Non-specific binding of non-mercurated ^{3}H-labeled RNA to the column was negligible (Table I).

10. Sizing of In Vitro Synthesized RNA on Sucrose Gradients

In vitro synthesized RNA was sized on 15-30% sucrose gradients in SW41 tubes. RNase free sucrose (Schwarz/Mann) was used and buffered with 10 mM Tris, pH 7.4, 0.1 M NaCl, 1 mM EDTA, 0.5% SDS. Gradients were run 15 hr at 25K at 24°C. 0.5 ml fractions were collected and assayed for TCA precipitable radioactivity.

11. Preparation of Cellular RNA's

Nuclear and cytoplasmic RNA were prepared as described elsewhere (Levy W. and McCarthy, 1975, 1976; Getz et al., 1975). For a typical preparation, 1 liter of Drosophila cells or L-cells was collected by centrifugation at 1500xg for 10 min. All subsequent steps until the phenol extractions were done at 0-4°C. The cells were washed several times in 10 mM Tris, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$ (RSB buffer) and then resuspended in 40 ml RSB containing 1% diethylpyrocarbonate. For Drosophila cells, 25 µg/ml polyvinyl sulfate and 10^{-3} M spermine was

also added. NP40 was added to 0.5% and the cells were lysed by homogenizing in a Dounce homogenizer. Nuclei were pelleted by centrifugation at 4000 rpm for 5 min in Sorvall HB4 rotor. The cytoplasmic supernatant was made 0.1 M in NaCl, 0.01 M in EDTA and 0.5% in sodium dodecyl sulfate, and extracted as described below. The nuclear pellet was washed several times in RSB and then resuspended in 40 ml 0.1 M NaCl, 10 mM Tris, pH 7.4, 10 mM EDTA, 0.5% SDS. Nuclear or cytoplasmic RNA was extracted by adding to each aqueous phase 1 volume phenol-chloroform equilibrated in 10 mM Tris, pH 7.4, 10 mM NaCl, 10 mM EDTA and shaking for 10 min at room temperature. The phases were separated by 10 min centrifugation at 6K in a Beckman JS 7.5 rotor. The interphase and aqueous phase was re-extracted several times by adding 1 volume chloroform containing 4% isoamyl alcohol. The aqueous phase was removed and precipitated with 2 volumes of ethanol. The precipitated nucleic acids were recovered by centrifugation at 6K for 60 min at -10°C in a Beckman JS 7.5 rotor. In the case of nuclear RNA, the precipitate was freed of DNA by treating with 50 µg/ml DNase I in 10 mM Tris, pH 7.4, 0.1 M NaCl, 5 mM $MgCl_2$ for 2 min at room temperature while vortexing, followed by 20 min at 37°C. SDS was added to 0.5% and the mixture was extracted as before and precipitated in ethanol. The RNA was recovered by centrifugation, resuspended in 10 mM Tris, pH 7.4, 10 mM NaCl, 1 mM EDTA, 0.5% SDS and passed over a Sephadex G-100 column equilibrated in the same. The material eluting in the void volume was pooled and precipitated in ETOH. This designated total nuclear RNA.

Total cellular RNA from L-cells was prepared by collecting cells by centrifugation, lysing them in 10 mM Tris, pH 7.4, 10 mM NaCl, 10 mM EDTA, 1% diethylpyrocarbonate, and 0.5% SDS and following the extraction

procedure given for nuclear RNA.

12. Preparation of Cellular DNA for Hybridization

Nuclei were prepared as described for chromatin preparation and resuspended in 10 mM Tris, pH 7.5, 1xSSC (1xSSC = 0.15 M NaCl, 0.015 M NaCitrate) + 0.5% SDS. DNA was extracted by adding 1 volume phenol-chloroform equilibrated in 10 mM Tris, pH 7.5, 1xSSC, shaking at room temperature for 15 min. The phases were separated by centrifugation and then the interphase and aqueous phase were re-extracted several times with phenol-chloroform-isoamyl alcohol (1:1:0.04). The DNA was precipitated with ethanol spooled out, resuspended in 0.1 M NaAc, pH 4.2 and depurinated at 37°C for 30 min. The pH was then brought to greater than 13 by adding NaOH and the mixture was hydrolyzed at 50°C for 20 min. The mixture was neutralized with HCl and ethanol precipitated. The precipitate was recovered by centrifugation, resuspended in H_2O, and dialyzed against several hundred volumes H_2O, with several changes of H_2O.

13. Preparation of Radioactive Probes for Hybridizations

a. Mouse satellite DNA was donated by Dr. Gerhart Ryffel and was prepared as described (Ryffel et al., 1975).

b. cDNA's to poly(A)-containing nuclear RNA was prepared as described (Levy W. and McCarthy, 1975) and was donated by Dr. Beatriz Levy W. For fractionated cDNA, aliquots of cDNA were annealed to nuclear RNA to a $R_o t$ of 12 M sec and the double stranded cDNA fraction (representing frequent sequences) was separated from the single stranded cDNA fraction (representing infrequent sequences) by adsorption chromatography on hydroxylapatite as described (Levy W. and McCarthy, 1975).

c. Nuclear RNA from mid log phase Drosophila cells was prepared by concentrating the cells 10-fold into fresh medium and labeling 20 min

with 10 μCi/ml ^{3}H-uridine (28 Ci/mmole; Schwarz/Mann). RNA was then extracted as described above for total nuclear RNA.

14. Hybridizations

Hybridizations of the in vitro transcripts with the total or fractionated cDNA probe made from poly(A)-containing nuclear RNA were carried out in 0.24 M equimolar phosphate buffer containing 1 mM EDTA and 0.01% SDS. 500 to 1000 cpm of cDNA and the RNA or DNA were sealed in 5 μl capillaries, boiled for 10 min and incubated at 70°C for the desired periods of time. Hybridizations of L-cell transcripts with satellite heavy strand DNA were carried out in 0.3 M NaCl, 10 mM Tris, pH 7.4 with about 500 cpm of heavy strand (2×10^{5} cpm/μg). A small amount (20 μg/ml) of calf thymus or bacterial DNA was added as carrier. Chromatin transcripts were present in at least 1000-fold excess over satellite heavy strand and DNA transcripts were present in at least 100-fold excess over satellite DNA. Reaction volumes were either 0.5 ml in sealed tubes or 50 μl in sealed capillaries. Samples were boiled for 10 min and incubated at 60°C for various lengths of time. At the end of the incubation the samples were cooled on ice, diluted into 4 ml 0.3 M NaCl, 3 mM $ZnCl_2$, 30 mM NaAc, pH 4.5, and challenged with S1 nuclease as described by Leong et al. (1972). The extent of hydrolysis of single or double stranded DNA at 5 μg/ml in 2 hr in the presence of various levels of S1 is shown in Figure 1. Background levels of digestion of 10-20% were routinely obtained when S1 was used at 600-1000 units/ml and when DNA concentrations were greater than or equal to 5 μg/ml. This level was subtracted from the experimental points.

Hybridizations of unlabeled in vitro transcribed RNA with labeled total nuclear RNA were performed in 6xSSC, 45% formamide at 50°C as

Figure 1. Kinetics of digestion of single and double stranded ^{3}H-thymidine labeled DNA (5 µg/ml) with various amounts of S1 nuclease.

(O---O): single stranded DNA + 100 units/ml S1 nuclease

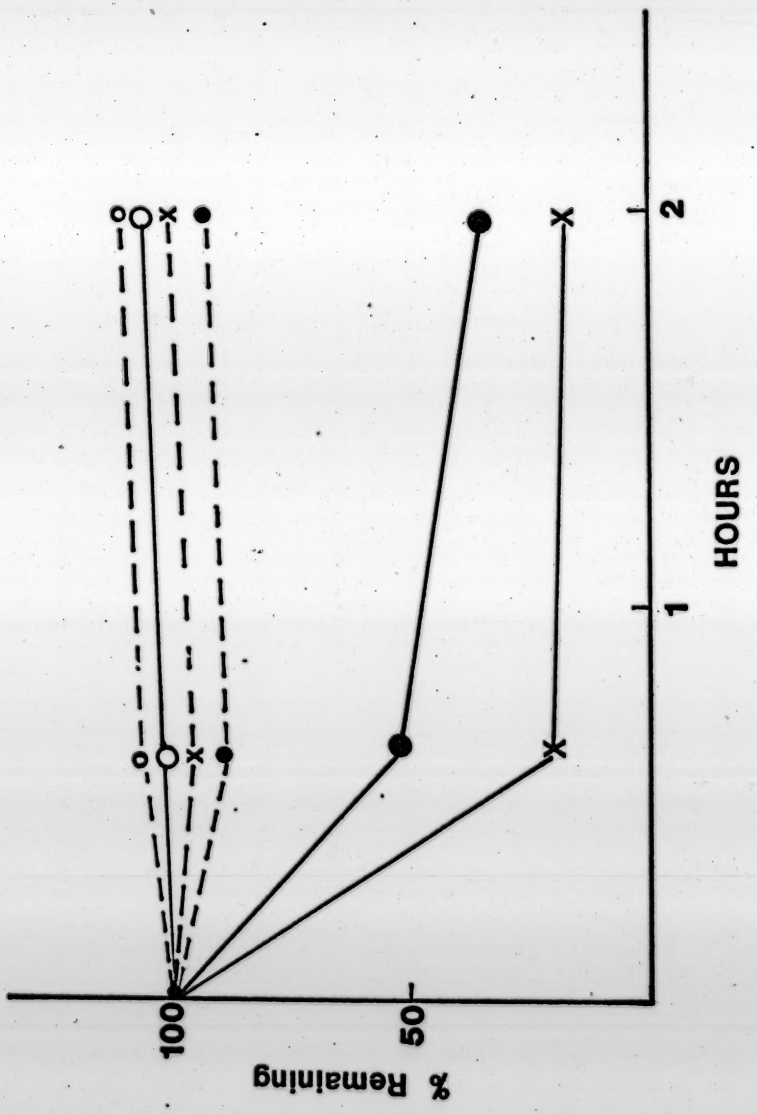
(O——O): single stranded DNA + 10 units/ml S1 nuclease

(X---X): single stranded DNA + 1000 units/ml S1 nuclease

(●---●): double stranded DNA + 10 units/ml S1 nuclease

(●——●): double stranded DNA + 100 units/ml S1 nuclease

(X——X): double stranded DNA + 1000 units/ml S1 nuclease


% Remaining
100
50
1
2
HOURS

described (Obinata et al., 1975) except that 0.01% SDS was added. After the desired incubation time the samples were diluted into 1 ml 2xSSC and incubated with 5 µg/ml pancreatic RNase for 30 min at 37°C. TCA precipitable material was collected on glass fiber filters (Whatman, GF/C) and washed with cold 5% TCA + 10 mM sodium pyrophosphate and ethanol. Filters were counted in Omnifluor (New England Nuclear) solution in a Beckman liquid scintillation spectrometer.

C. RESULTS

I. Analysis of L-Cell Chromatin Transcription under Low Salt Conditions

A. Preparation of Chromatin

Chromatin was prepared for these experiments following the method of Bonner et al. (1968). For solubilization, chromatin was sheared at 3000 psi in a French pressure cell, a treatment which produces fragments of chromatin containing DNA of double stranded molecular weight up to about 3×10^6 daltons, or about 5000 base pairs (McCarthy et al., 1973). Under the conditions of the preparation, nuclei were never exposed to divalent cations and chromatin was never exposed to ionic strengths higher than 10 mM Tris, pH 8. Such conditions tend to minimize protein sliding or rearrangement. Chromatin was reconstituted following the procedure of Kleiman and Huang (1972) as described in Materials and Methods. High pH (pH 8) was used to enhance chromatin solubility, since these authors report that solubility is poor in 0.2 M NaCl when the pH is less than 7.5. I omitted the use of $NaHSO_3$, reported to inhibit histone proteases (Kleiman and Huang, 1972), since no histone degradation was observed on polyacrylamide gels for mouse L-cell histones under the conditions of reconstitution (data not shown). During reconstitution, histones reassociate with DNA as the salt concentration is reduced to 0.1 M in the presence of 5 M urea. Most of the non-histone proteins return during dialysis from 0.1 M NaCl, 5 M urea, to dilute buffer (10 mM Tris, pH 8) (Chae, 1975).

B. Transcription of Chromatin

Chromatin was transcribed under conditions described by Axel et al. (1975) which attempt to minimize perturbations of the system by using

low ionic strength (10 mM Tris-Cl, pH 7.9, 1 mM $MnCl_2$). The experiments shown in Figures 2 and 3 characterize this transcription system. For these experiments, a small amount of ^{3}H-CTP (5 μCi/ml) was added to monitor transcription. Figure 2 shows the amount of RNA synthesized from chromatin at varying enzyme/template ratios. Maximal synthesis was obtained at a mg enzyme/mg template ratio of about 1 so this level of enzyme was chosen for subsequent experiments. This level is comparable to levels used by other investigators (Astrin, 1975; Axel et al., 1975). Figure 3 shows the kinetics of transcription using this level of enzyme. Synthesis continued for about 45 min and produced about 60 μg RNA per mg template chromatin.

C. Isolation of Transcripts

After synthesis for 45 min, the transcript was isolated as described in Materials and Methods, using two cycles of DNase I treatment in order to free the RNA from DNA. To demonstrate the absence of contaminating DNA in the RNA, a portion of the transcript was brought to pH >13 with NaOH (final NaOH concentration, 0.3 M), and boiled for 10 min. The samples were then neutralized with HCl and adjusted to the proper salt and buffer conditions for hybridization.

Since chromatin contains an endogenous RNA component whose amount varies from preparation to preparation, it was necessary to determine the content of endogenous RNA in the chromatin undergoing transcription and make the appropriate corrections for the amount of transcript synthesized in vitro by E. coli RNA polymerase. A sample of chromatin without added enzyme was incubated and extracted in parallel with a sample of chromatin with added enzyme. In most preparations, 1 mg of chromatin transcribed in vitro yielded 60 μg of RNA after isolation, of

Figure 2. *In vitro* transcription of mouse L-cell chromatin with increasing amounts of *E. coli* RNA polymerase. After 30 minutes of incubation, the reaction mixture containing ^{3}H-UTP was TCA precipitated and the radioactivity was determined.

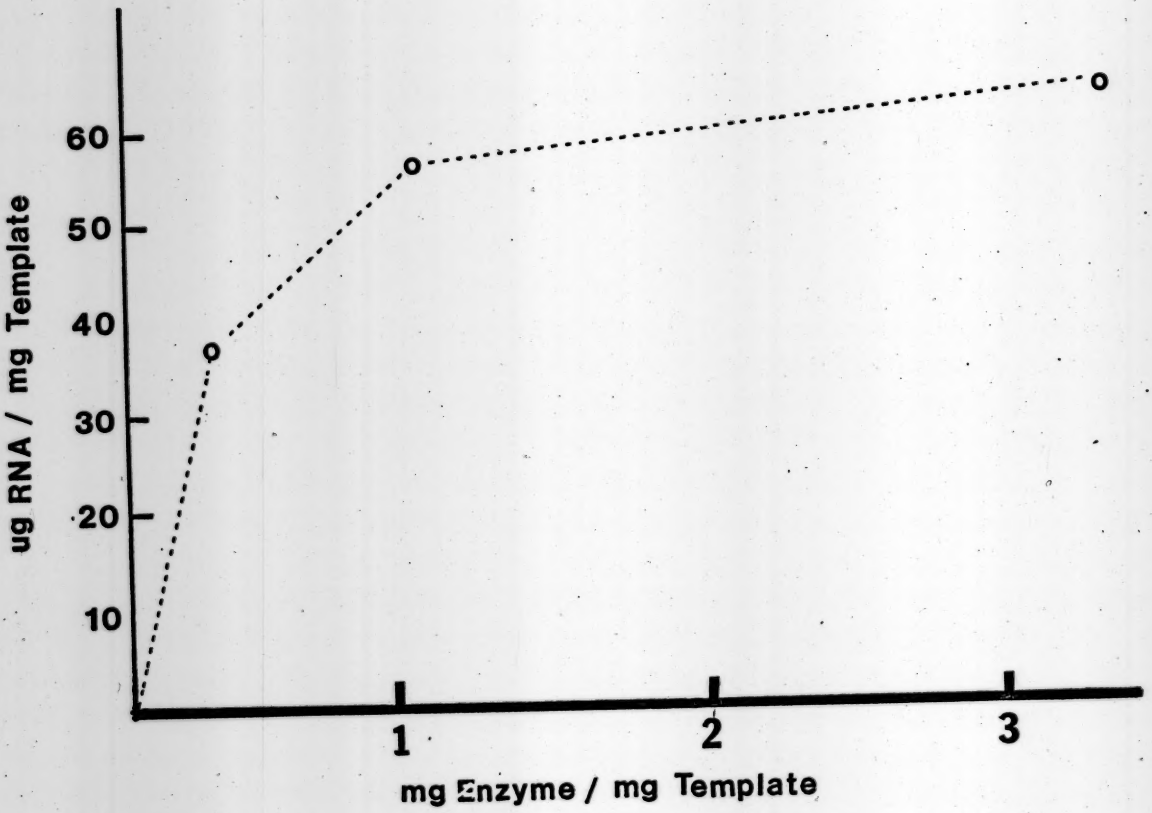
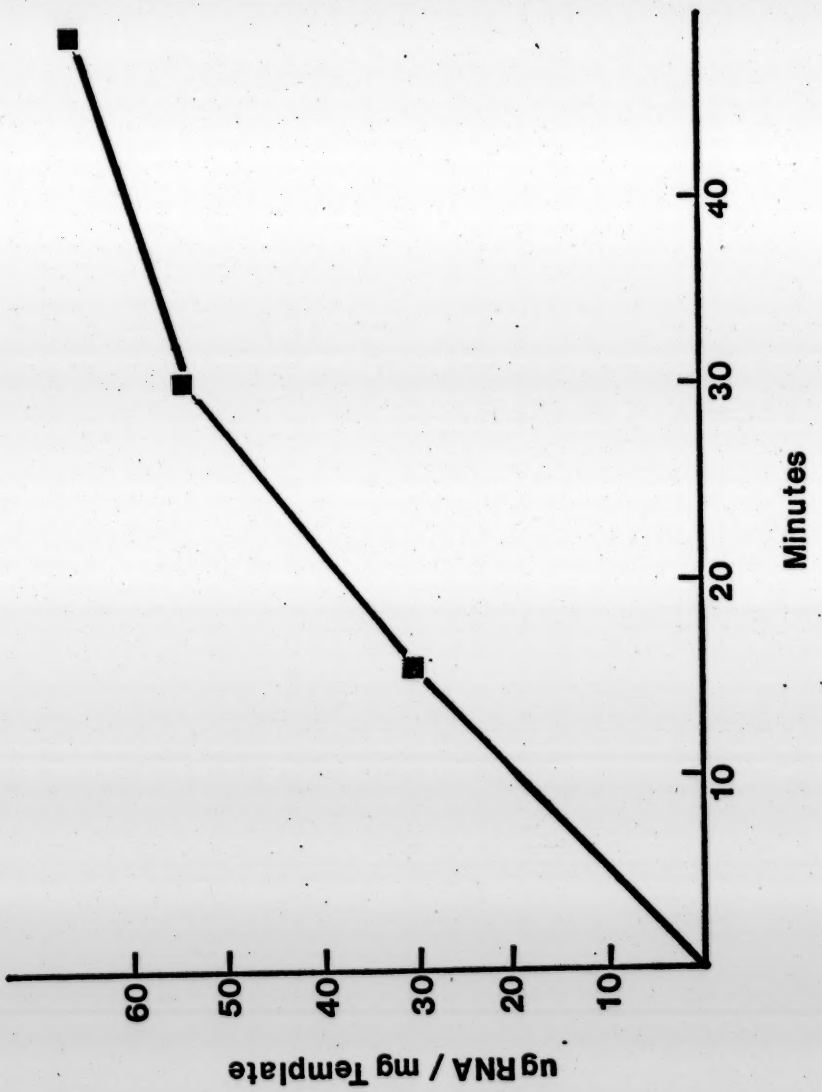


Figure 3. Kinetics of *in vitro* transcription from L-cell chromatin. The reaction mixture contained ^{3}H-UTP (5 μCi/ml) in addition to the usual reactants (see Materials and Methods). TCA precipitable counts incorporated were determined at various times after addition of RNA polymerase.


ug RNA / mg Template
10
20
30
40
50
60
10
20
30
40
Minutes

which 30 µg were present endogenously. For such a situation, the $C_o t$ curves needed to be corrected by a factor of 2. Since satellite sequences are absence from *in vivo* RNA (Flamm *et al.*, 1969), it was not necessary to consider correction of the data for endogenous RNA which would hybridize with the satellite DNA probe.

Endogenous RNA and total L-cell RNA were purified and used as controls for the *in vitro* transcriptions.

D. Hybridizations

Hybridizations between L-cell satellite DNA and various RNA preparations were performed as described in Materials and Methods and the percent hybrid was determined by digestion with S1 nuclease. Figure 4 shows the hybridization kinetics between satellite heavy strand and a 100-fold excess of total L-cell DNA or DNA transcript. Taking the level of satellite sequences in DNA to be 10% and allowing for the finding that DNA-DNA reactions proceed with 2-fold faster kinetics than do DNA-RNA reactions (Melli *et al.*, 1971) the level of satellite sequences in the DNA transcript is also about 10%, indicating random transcription by *E. coli* RNA polymerase with respect to satellite. This is consistent with the results of Maio and Kurmit (1974). Also shown in Figure 4 are hybridization curves for endogenous RNA and total L-cell RNA. From the saturation values obtained, and taking the excess of RNA to probe to be 1000, we calculate that satellite sequences comprise less than one part in 10^4 in the endogenous RNA and less than one part in 10^5 in total L-cell RNA.

Figure 5 shows the hybridization kinetics between satellite H-strand and a 1000-fold excess of chromatin transcript or reconstituted chromatin transcript. The DNA transcript curve is included as comparison.

Figure 4. Hybridization kinetics of ^{3}H-thymidine labeled satellite heavy strand (ordinate) with total L-cell DNA (X——X), DNA transcript (Δ---Δ), endogenous RNA (●——●) and total L-cell RNA (▫——▫).

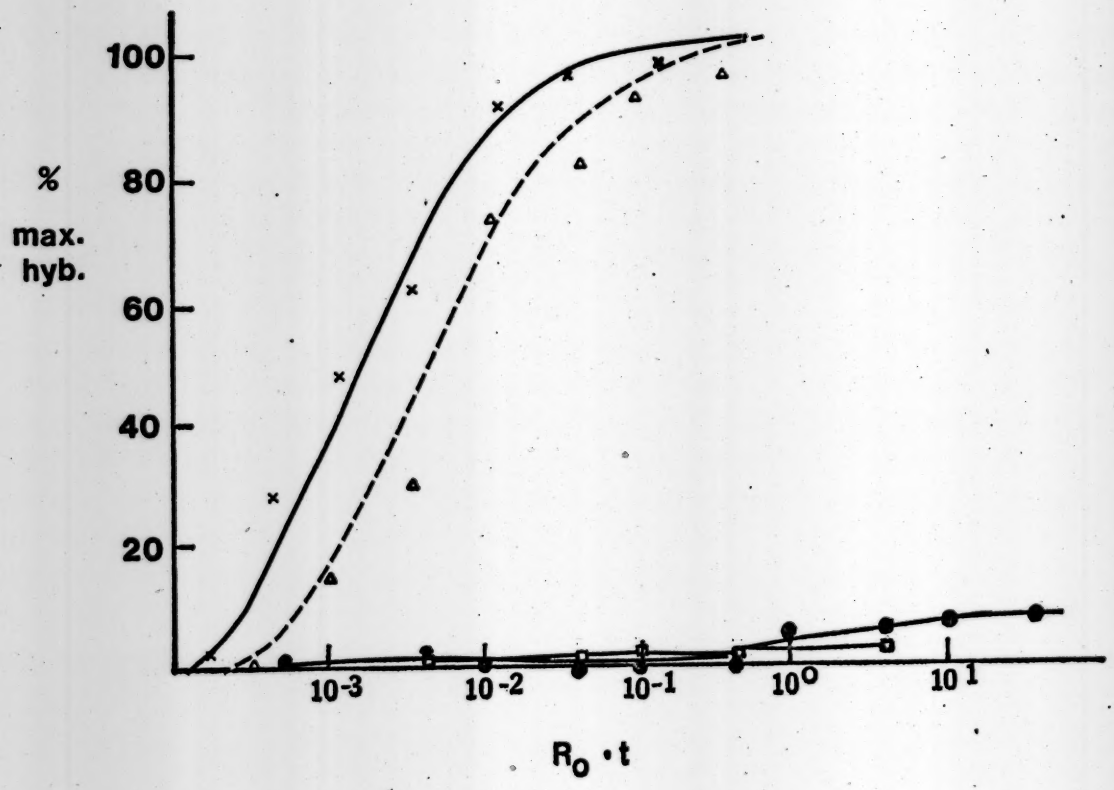
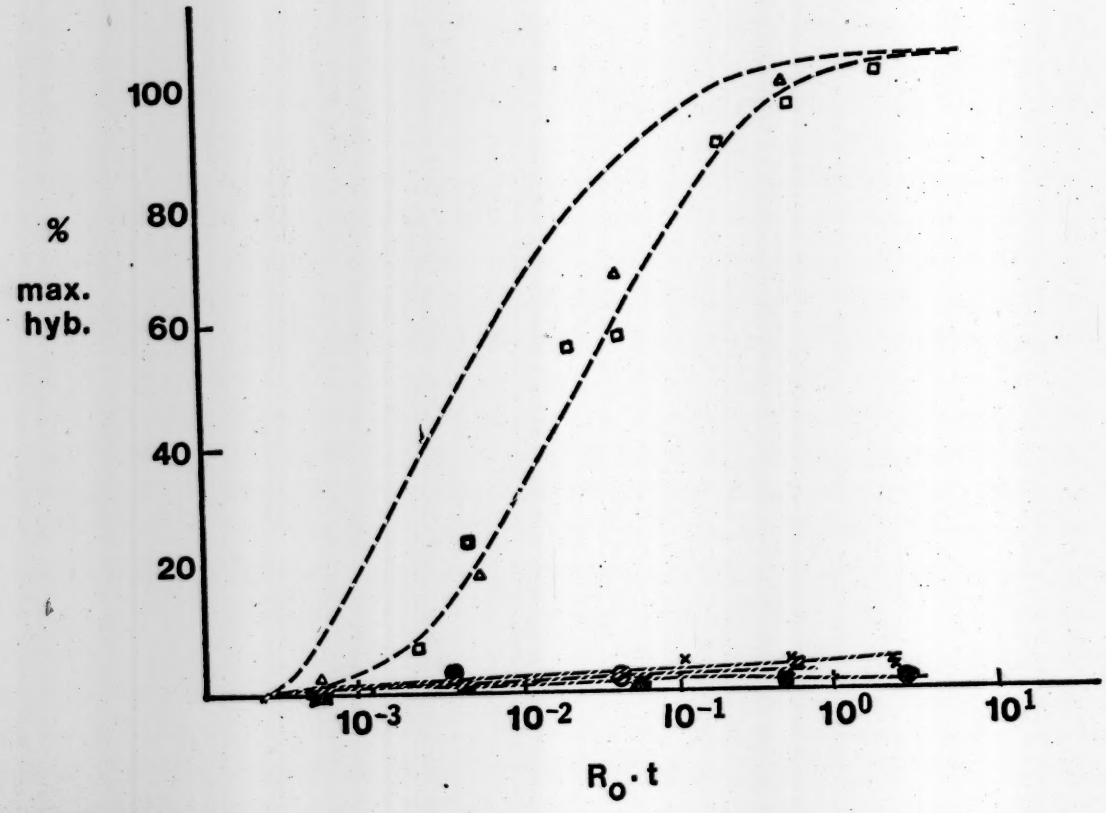


Figure 5. Hybridization kinetics of ^{3}H-thymidine labeled satellite heavy strand (ordinate) with chromatin transcript (Δ---Δ), reconstituted chromatin transcript (◘ --- ◘), alkali treated DNA transcript (X---X), alkali treated chromatin transcript (■ --- ■), alkali treated reconstituted chromatin transcript (●---●). Reference curve from Fig. 8 for DNA transcript hybridization kinetics is also included (---).



The kinetics with which the chromatin transcript hybridizes are about 4-fold slower than those of the DNA transcript, indicating that satellite sequences are preferentially repressed in chromatin. However, a significant level of expression, about 2%, is still observed. Reconstituted chromatin transcript hybridizes with similar kinetics as does that from native chromatin, indicating that the dissociation and reassociation process has not affected the degree to which satellite sequences are expressed. Also included in Figure 5 are curves obtained after treatment of the transcripts with alkali. This treatment abolishes all hybridization indicating that the transcripts were free of DNA.

II. Transcription of Chromatin in the Presence of Hg-UTP

A. Discussion of Problem of endogenous RNA.

In all chromatin preparations there is a variable amount of a rather undefined population of RNA molecules termed endogenous RNA, consisting of pieces of residual *in vivo* transcripts. When one is monitoring *in vitro* transcription by assaying for the presence of sequences which are not transcribed *in vivo*, as was done above, it is possible to deal with the problem simply by determining the endogenous RNA content in the chromatin and making the appropriate corrections in the hybridization kinetics. However, in the experiments discussed below, I wished to assay for sequences which are transcribed *in vivo*. In such a case, the presence of endogenous RNA in the *in vitro* transcript preparation will obscure the hybridization kinetics and make interpretations difficult. Therefore, in the following experiments, a mercurated precursor, 5-mercuriuridine triphosphate (Hg-UTP) was incorporated into the *in vitro* transcript, making it possible to purify the transcript from endogenous RNA. In the same way, template DNA is removed, eliminating the need to digest the transcript with DNase I.

B. Preparation of 5-Mercuriuridine Triphosphate (Hg-UTP)

Dale and Ward (1975) have reported that pyrimidine derivatives are quantitatively mercurated in 3 hours at 50°C in buffered solutions of NaAc containing mercuric acetate. The reaction proceeds via electrophilic substitution in which mercury displaces a proton at the C-5 position ($R\text{-}H + HgX_2 \rightarrow\rightarrow RHgX + HX$). Reactions performed in this manner were chilled on ice and passed over a Sephadex G-10 column to remove unreacted mercury. Figure 6 shows a typical elution profile from a 100 ml column of a 20 ml reaction mixture containing 100 mg UTP. Hg-UTP was monitored by OD_{260} and the presence of free mercury ions was monitored by adding a few drops of 2 M NaOH to a small aliquot of the eluate. The formation of a yellowish precipitate (mercuric oxide) indicated the presence of Hg ions. The fractions containing Hg-UTP were pooled and diluted 5-fold with water. Further removal of unreacted mercury ions was accomplished by chromatography on a DEAE-cellulose column equilibrated in 0.01 M triethylaminebicarbonate ($TEA\text{-}HCO_3$). The material which did not wash off the column was batch eluted with 0.5 M $TEA\text{-}HCO_3$. This material was flash evaporated twice with methanol, evaporating each time. The Hg-UTP was finally resuspended in 10 mM Tris, pH 7.5 at about 5 mg/ml.

The mercurated product was characterized by its spectral properties and by thin layer chromatography on polyethyleneimine cellulose (Brinkman Instruments). The absorption maxima of mercurated pyrimidine compounds occur at longer wavelengths (by 5 nm) than the parent compounds, and in the case of Hg-U compounds, the entire spectrum is red shifted (Dale and Ward, 1975). Figure 7 shows spectra of 50 µg/ml solutions of UTP and the mercurated product, and demonstrates that mercuration has occurred.

Chromatography on polyethyleneimine cellulose with increasing

Figure 6. Sephadex G-10 elution profile of Hg-UTP. 100 mg Hg-UTP in 20 ml 0.1 M $Hg(Ac)_2$, 0.1 M NaAc, pH 6.0, was passed over a 100 ml G-10 column equilibrated in 0.1 M NaAc, pH 6.0. Hg-UTP was monitored by OD_{260} and the presence of Hg ions was determined by adding a few drops of 2 M NaOH to an aliquot of each fraction and checking for the formation of a yellowish precipitate. Fractions containing Hg are indicated by hatch marks. Fractions containing Hg-UTP (fractions 11-28) were pooled.

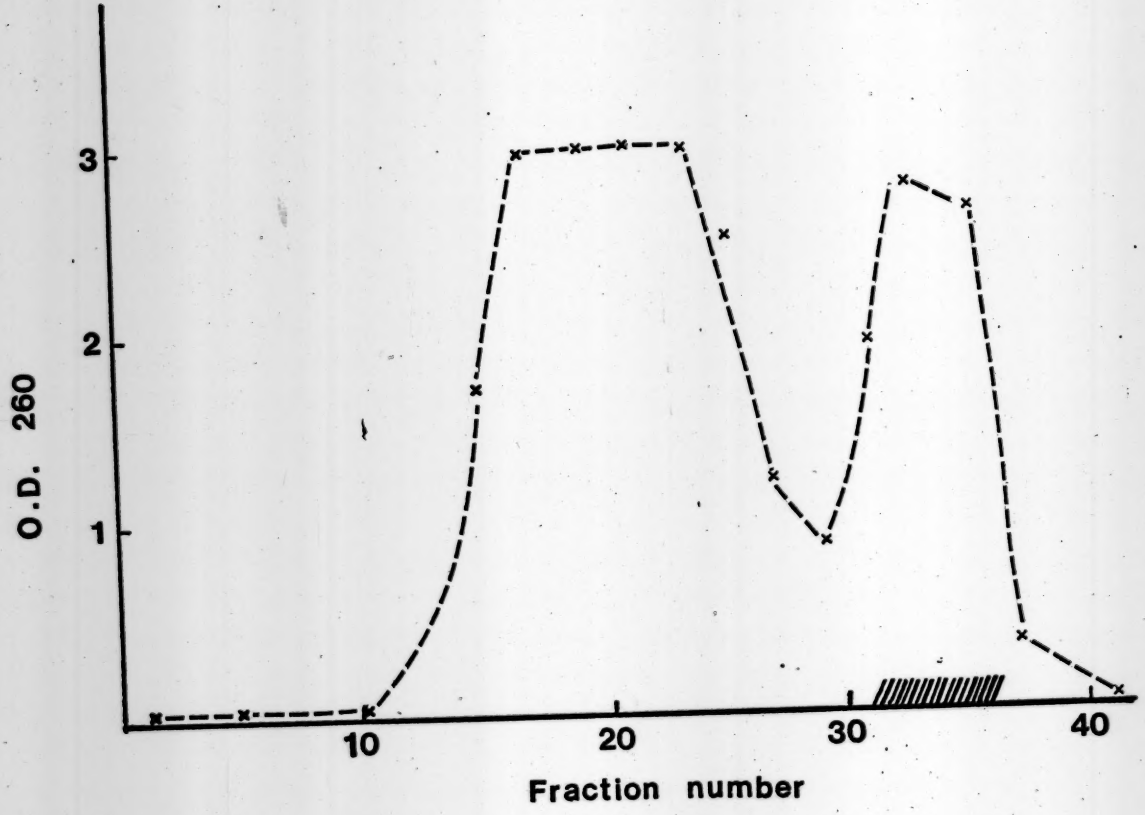
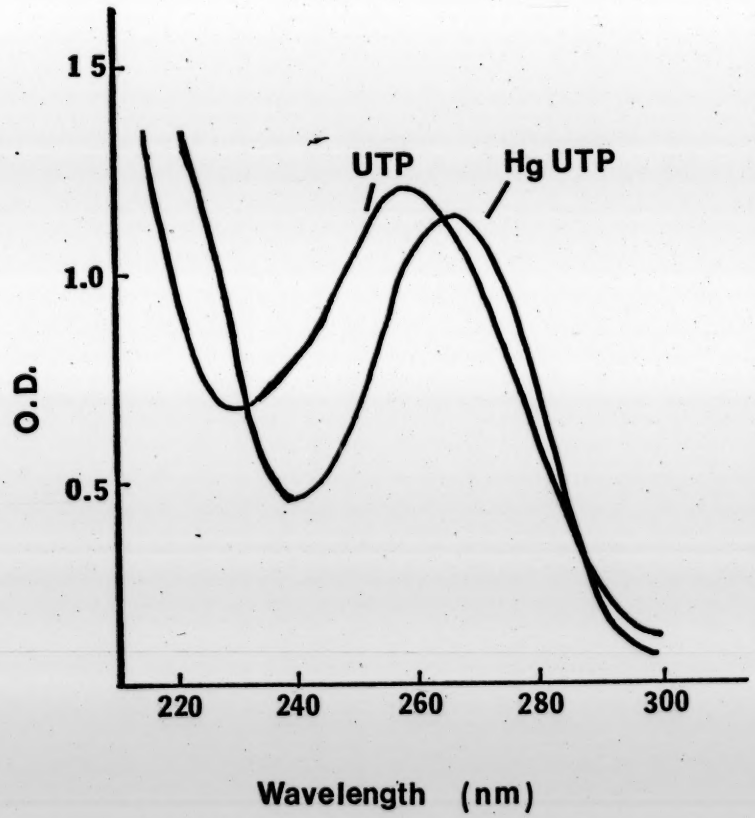


Figure 7. Spectra of UTP and Hg-UTP solutions (50 µg/ml in 10 mM Tris, pH 7.5).


1.5
1.0
0.5
O.D.
UTP
Hg UTP
220
240
260
280
300
Wavelength (nm)

concentrations of LiCl (Randerath and Randerath, 1964) enables one to distinguish the mercurated compound from the non-mercurated compound. Mercurated UTP does not migrate significantly from the origin under these conditions. The addition of mercaptoethanol to the sample increases the electrophoretic mobility to 85-90% of that of the non-mercurated compound. When the mercurated product was analyzed in this way, it was apparent that virtually all of the UTP had been mercurated, and that none of it was degraded to the di- or monophosphate form.

C. Preparation of Chromatin

Chromatin was prepared from nuclei which had been purified in the presence of sucrose and divalent cations as stabilizers. Again, attempts were made to treat the chromatin as gently as possible so as to minimize perturbations of the system. To lyse the nuclei and solubilize the chromatin, the nuclear pellet was resuspended in 10 mM Tris, pH 8 and sheared at very low pressures (1000 psi) in a French pressure cell. This produced chromatin fragments with DNA of double stranded molecular weight up to about 14×10^6 daltons (25,000 base pairs) as shown in Figure 8.

D. Transcription of Chromatin

Chromatin was transcribed under conditions more closely resembling the physiological state (Astrin, 1975), that is, salt concentrations were set at 150 mM. Under these conditions, synthesis continued for 4 hr and produced large amounts of *in vitro* transcript. Figure 9 shows the amount of RNA synthesized in 3 hr from Drosophila DNA or chromatin at increasing polymerase/ template ratios. No Hg-UTP was added in this experiment and 5 μCi/ml of ^{3}H-CTP (23 Ci/mmole) was added as a tracer. The data show that DNA is transcribed with higher efficiency than is

Figure 8. Size distribution of DNA in chromatin (upper curve). Scan of a photograph taken from a 0.9% agarose gel stained with ethidium bromide. Eco R1 endonuclease fragments of lambda phage DNA were used as size markers (lower curve).

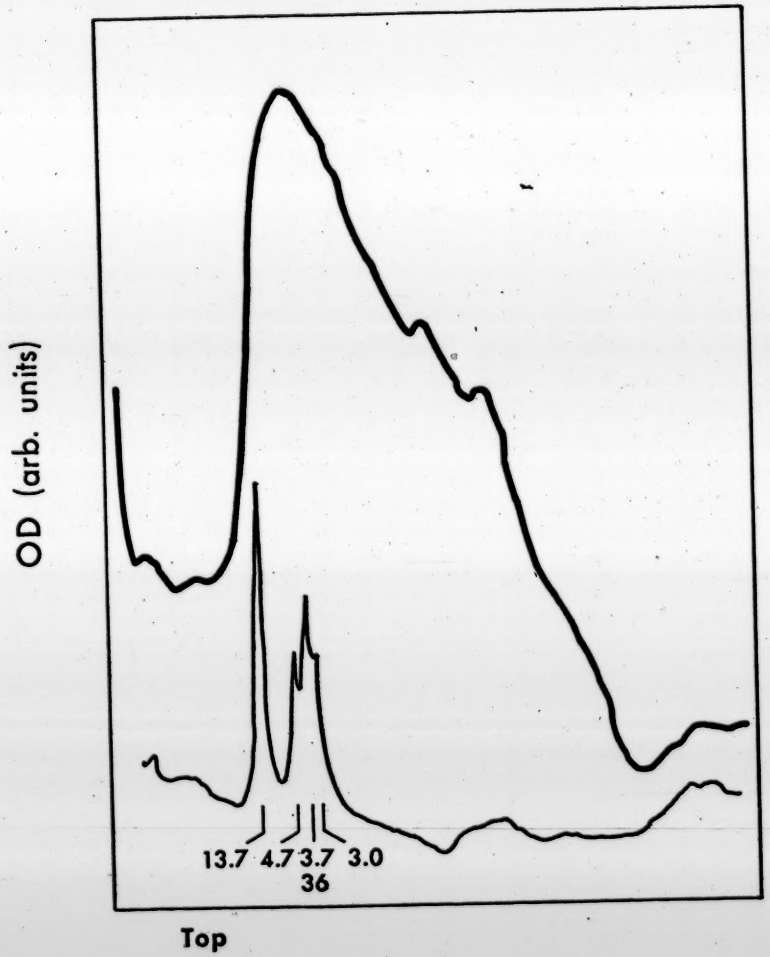
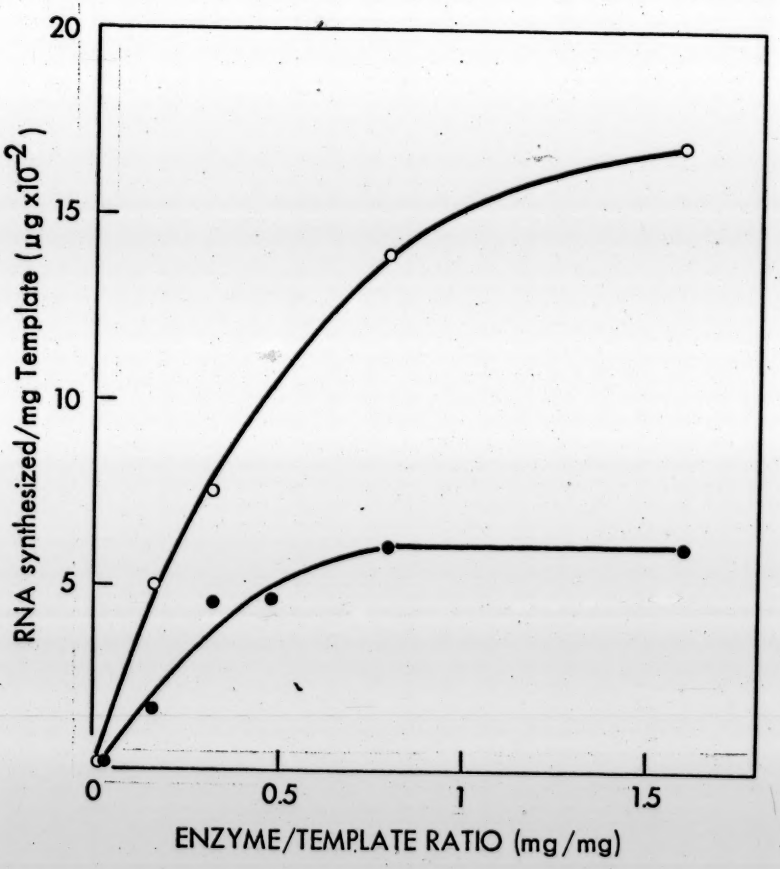

OD (arb. units)
13.7
4.7
3.7
3.0
36
Top

Figure 9. In vitro transcription of chromatin (●) and DNA (○) at 25°C with increasing amounts of E. coli RNA polymerase. After three hours of incubation, the reaction mixture (containing 5 μCi/ml ^{3}H-CTP) was precipitated with 5% TCA and the precipitable radioactivity was determined.


20
15
10
5
0
0.5
1
1.5
RNA synthesized/mg Template (μg x10⁻²)
ENZYME/TEMPLATE RATIO (mg/mg)

chromatin and that maximal transcription is achieved at a mg polymerase/template ratio of 0.8 for chromatin. This was therefore the level of enzyme used for further experiments. The in vitro transcription kinetics from Drosophila DNA and chromatin using polymerase/template ratios of 0.8 and 4 mM ATP, GTP, CTP, 1 mM UTP and 5 μCi/ml ^{3}H-CTP are illustrated in Figure 10. Under these conditions the reaction proceeded for 4 hr and yielded about 0.6 mg RNA for each mg chromatin and more than twice as much from DNA. It appears from the length of time during which the reaction occurs, that chain reinitiation is occurring in vitro. An experiment in which chromatin was incubated in a reaction mixture in the absence of E. coli RNA polymerase, showed that the endogenous enzyme remained active for about 30 min and synthesized about 30 μg RNA per mg chromatin, a negligible amount relative to that synthesized in the presence of the E. coli enzyme.

Substitution of Hg-UTP for UTP in the reaction mixture reduced the extent of RNA synthesis with either DNA or chromatin as template. This result is in agreement with the results of Dale and Ward (1975) who reported a similar effect using polyd(A-T) as a template. The amount of RNA synthesized after 4 hr with Hg-UTP as a precursor was about 0.4 to 0.5 mg per mg template in the case of chromatin and about 1 mg per mg template in the case of DNA.

The kinetics of transcription with L-cell chromatin as template is shown in Figure 11. The reaction mixture contained 1 mM Hg-UTP and 5 μCi/ml ^{3}H-CTP. About 0.5 mg RNA was synthesized from L-cell chromatin in 4 hr when transcription was performed either at 37°C or 25°C. Unsheared chromatin was liberated from nuclei by homogenizing gently in a Dounce homogenizer in 10 mM Tris, pH 8, and also transcribed. In this case

Figure 10 Kinetics of in vitro transcription at 25°C from Drosophila chromatin or DNA. ATP, GTP, and CTP were 4 mM, 5 μCi/ml ^{3}H-CTP. DNA template, 1 mM UTP (▲); DNA template, 1 mM Hg-UTP (Δ); chromatin template, 1 mM UTP (●); chromatin template, 1 mM Hg-UTP (○).

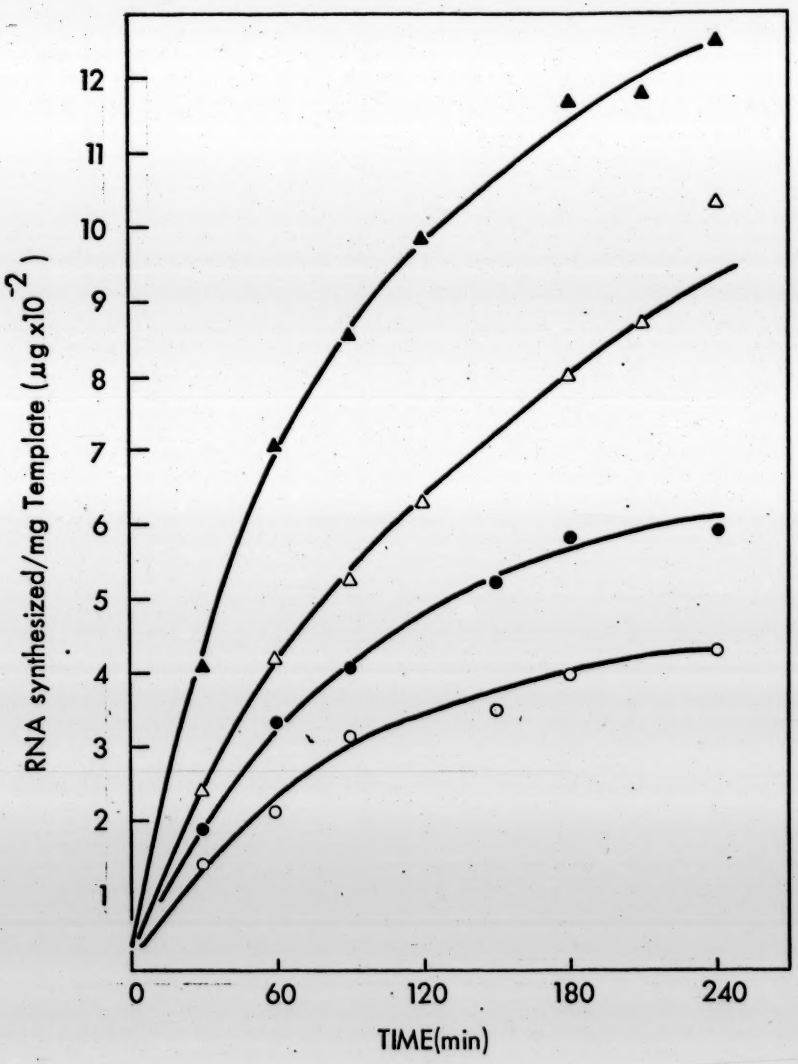
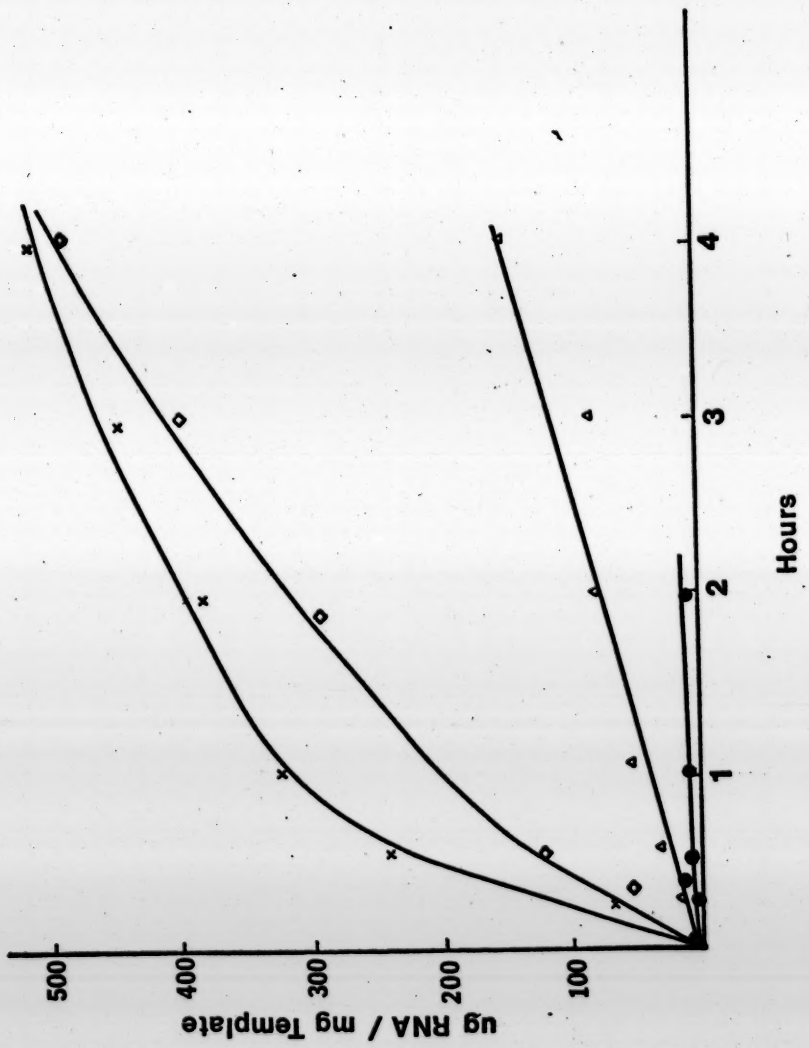

RNA synthesized/mg Template (µg x10^{-2})
12
11
10
9
8
7
6
5
4
3
2
1
0
60
120
180
240
TIME(min)

Figure 11. Kinetics of *in vitro* transcription from L-cell chromatin. ATP, GTP and CTP were 4 mM. Hg-UTP was 1 mM. 5 μCi/ml ^{3}H-CTP was added. Chromatin template transcribed at 37°C (X); chromatin template transcribed at 25°C (◆); unsheared chromatin template transcribed at 37°C (Δ); endogenous enzyme activity from chromatin template at 25°C or 37°C (●).



transcription was reduced and produced 150 µg RNA per mg template in 4 hr. No endogenous polymerase activity could be detected either at 25°C or at 37°C. The nucleic acids extracted from the reaction mixture after 4 hr incubation were freed of residual triphosphates by loading on a 50 ml Sephadex G-50 column equilibrated in TNE (10 mM Tris, pH 7.3, 0.1 M NaCl, 1 mM EDTA). Two ml of sample were loaded. A typical elution pattern is shown in Figure 12. The fractions in the void volume were pooled and precipitated with ethanol. This precipitate was resuspended in 1 ml TNE and loaded onto a sulfhydryl-Sepharose 6B column as described in Materials and Methods. A typical elution profile for an in vitro transcript synthesized in the presence of Hg-UTP and ^{3}H-CTP is shown in Figure 13. The nucleic acids which did not contain mercury were washed off the column and appeared in the first few fractions. Some mercury containing material did not bind to the resin as could be shown by rechromatography. This may be due either to a slow displacement of the mercury bound 2-mercaptoethanol by the sulfhydryl groups of the resin or to an overloading effect. The portion of the RNA which did not bind to the column varied between 10 and 40%. The RNA which bound and which could be eluted with 2-mercaptoethanol was ethanol precipitated and applied to a 15-30% sucrose gradient for size determination. As shown in Figure 14, this RNA sedimented with a peak at about 5S although a major fraction was greater than 10S. Non-specific binding of unmercurated RNA was negligible (see Table I).

E. Hybridization of L-Cell Transcripts to Satellite DNA

Several hybridization experiments were performed using RNA transcribed from L-cell chromatin and satellite DNA as a probe. These experiments allow a direct comparison to be made between this set of condi-

Figure 12. G-50 elution profile of the in vitro synthesized transcript from L-cell chromatin. Transcript was synthesized from 0.5 mg chromatin in a 1 ml reaction containing 4 mM each of ATP, GTP and CTP, 1 mM Hg-UTP, and 5 μCi/ml ^{3}H-CTP. Nucleic acids were extracted and ethanol precipitated (see Materials and Methods). The precipitate was resuspended in 1 ml TNE and applied to a 50 ml Sephadex G-50 column. 2.5 ml fractions were collected. OD_{260} (■); TCA precipitable radioactivity (X).

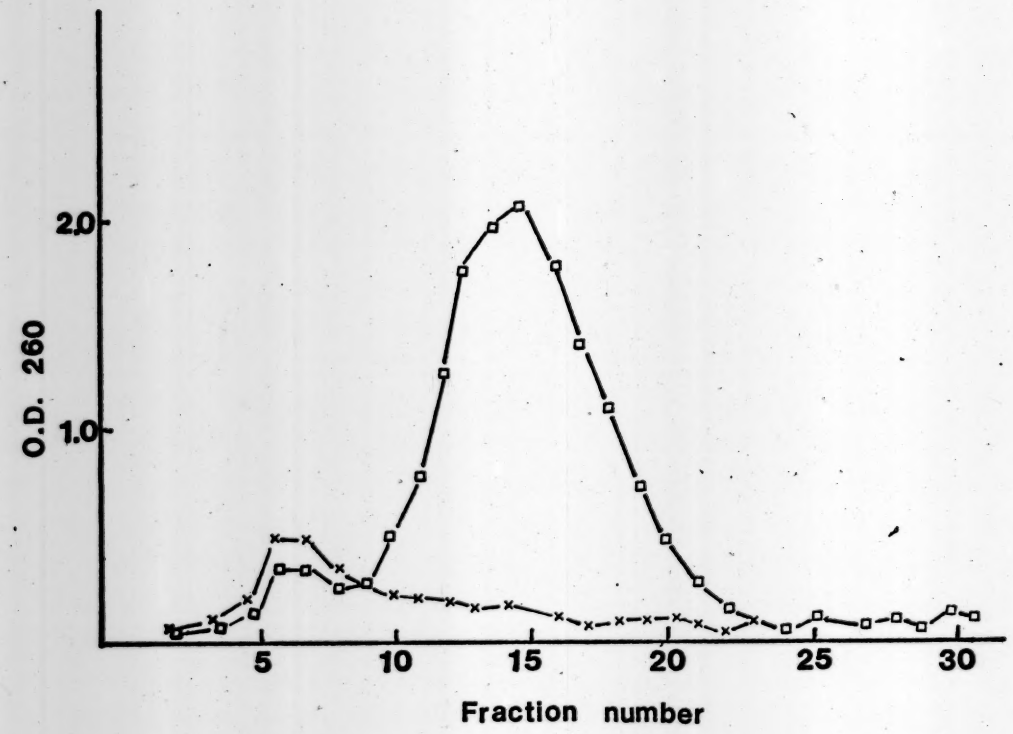


Figure 13. Chromatography of ^{3}H-CTP labeled *in vitro* synthesized RNA containing Hg-UTP on a sulfhydryl-Sepharose 6B column (3 ml). The arrow indicates the beginning of the elution with TNE + 0.1 M 2-mercaptoethanol. Fractions of 1 ml were collected and the TCA precipitable radioactivity was determined.

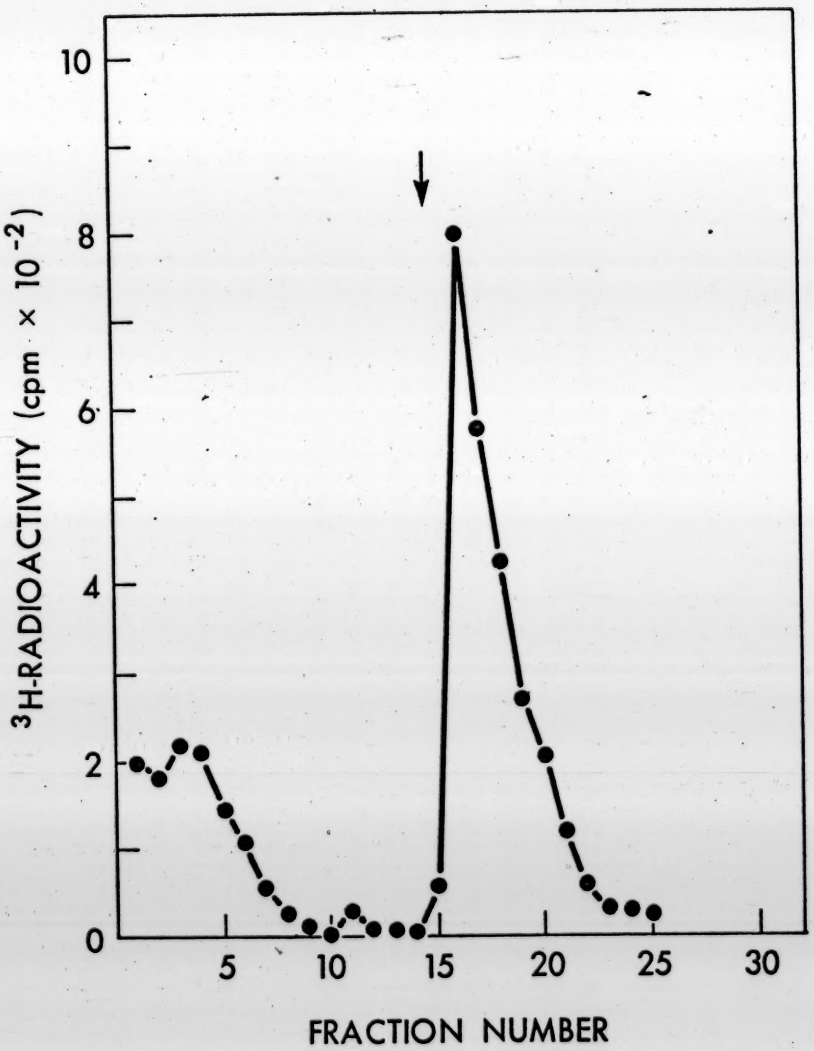

10
8
6
4
2
0
^{3}H-RADIOACTIVITY (cpm × 10^{-2})
5
10
15
20
25
30
FRACTION NUMBER

Figure 14. Size distribution of the *in vitro* synthesized RNA from chromatin after purification on sulfhydryl-Sepharose 6B. ^{3}H-CTP labeled RNA was loaded on a 15-30% sucrose gradient in 10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.5% SDS, in a SW40 rotor and centrifuged at 25,000 rpm for 15 hr at 24°C. Reading from left to right the arrows indicate the positions of 5S, 19S and 26S ribosomal RNA from Drosophila run in a parallel gradient.

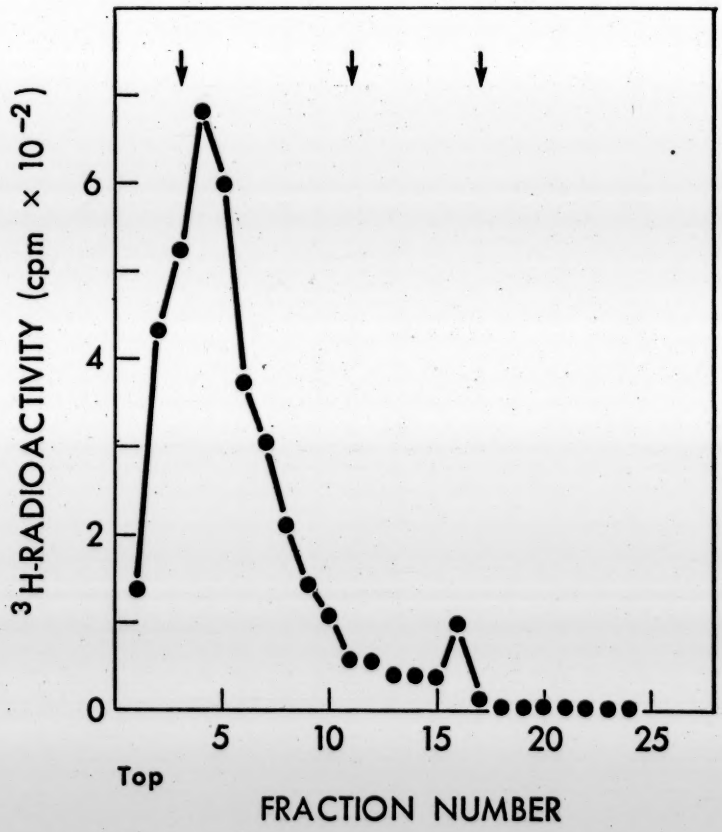

^{3}H-RADIOACTIVITY (cpm × 10^{-2})
0
2
4
6
Top
5
10
15
20
25
FRACTION NUMBER

TABLE I: Functional Test of Sulfhydryl-Sepharose 6B Column

Preparation applied to the column	^{3}H-Ribosomal RNA	^{3}H-non-polyadenylated RNA + unlabeled *in vitro* synthesized mercurated RNA 1:1 ratio	^{203}Hg
Percent not retained	94	91	2
Percent eluted with TNE + 2-mercaptoethanol	0.05	0.2	97.3
Percent remaining in column	Undetectable	undetectable	0.7

^{3}H-ribosomal RNA (8×10^{5} cpm), ^{3}H-polyA- nuclear RNA (43.8×10^{3} cpm) mixed in 1:1 ratio with unlabeled *in vitro* synthesized mercurated RNA or ^{203}Hg (New England Nuclear, 7.6×10^{5} cpm) in TNE were located on a 1 ml sulfhydryl-Sepharose 6B column and incubated for 30 min. The unbound material was washed off with 20 volumes TNE and the bound material was eluted with TNE + 0.1 M 2-mercaptoethanol. Total (in the case of ^{203}Hg) or TCA precipitable radioactivity was determined in the combined fractions.

tions for the isolation and transcription of chromatin and the earlier set of conditions described in part I. Figure 15 shows the hybridization kinetics of the labeled probe with the *in vitro* transcripts. Again, the DNA transcript hybridizes with kinetics which suggest that transcription was random. The transcript from chromatin hybridized with slower kinetics, being displaced some 5-fold to higher $R_o t$ values. Therefore, satellite sequences represent about 2% of the *in vitro* transcript. Figure 15 also shows that transcripts synthesized at 37°C or 25°C contain approximately the same amount of satellite sequences.

Other possible sources of artifact are the shearing process to which chromatin is subjected, and the presence of Mn^{++}, which has been reported to cause incorrect initiations of yeast polymerase I on ribosomal DNA (van Keulen *et al.*, 1975; Hager *et al.*, 1976). Therefore, it was interesting to determine if either of these treatments was responsible for the level of satellite transcription observed *in vitro*. Chromatin was prepared and transcribed as usual except that $MnCl_2$ was omitted from the transcription reaction mixture. Figure 16 shows that the kinetics of hybridization of transcript synthesized under these conditions is similar to those of transcript synthesized under the usual conditions, containing $MnCl_2$. Figure 16 also shows that transcript synthesized from unsheared chromatin, that is, chromatin produced by homogenizing the nuclei in 10 mM Tris, pH 8, produced a transcript which had similar amounts of satellite sequences. Therefore, it appears that although negligible amounts of satellite sequences are detectable in *in vivo* RNA's, there is a low but reproducible level detectable in *in vitro* transcripts produced under a variety of conditions. These results are summarized in Table II.

Figure 15. Hybridization kinetics of ^{3}H-thymidine labeled satellite heavy strand with L-cell DNA transcript (Δ), with chromatin transcript synthesized at 25°C (X), and with chromatin transcript synthesized at 37°C (◘).

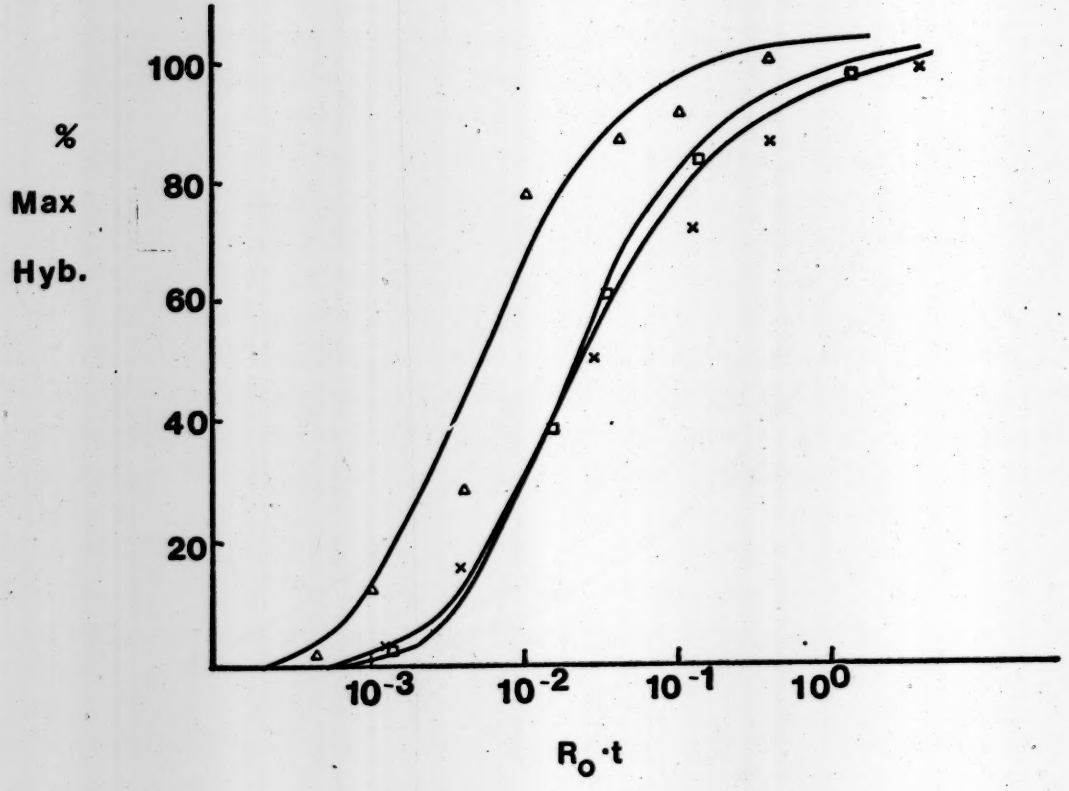


Figure 16. Kinetics of hybridization of satellite heavy strand with transcript from unsheared chromatin (◘), and with transcript synthesized in the absence of Mn^{++} (X).

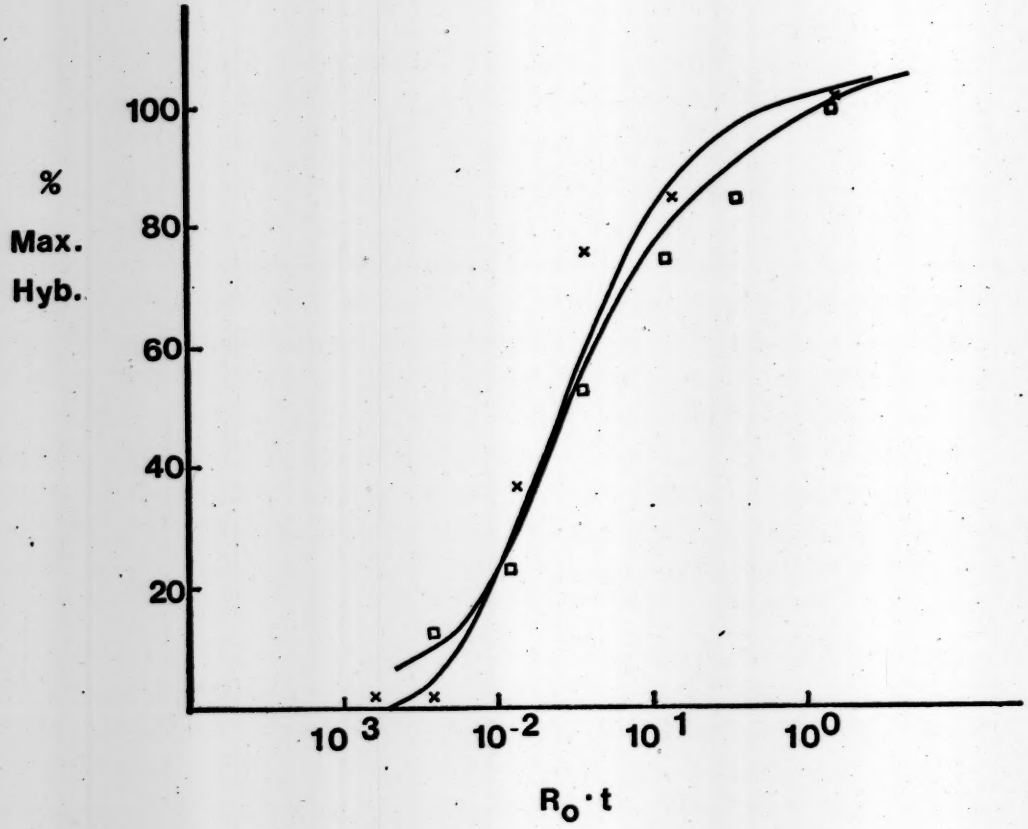


TABLE II: Amount of Satellite Sequences Found in Various RNA and DNA Preparations

DNA or RNA Preparation	Transcription conditions	% Satellite Sequences
1. Total DNA	--	10%
2. DNA transcript	Axel et al (1973) 37°C	10%
3. DNA transcript	Astrin (1975) + Hg-UTP 37°C	10%
4. Chromatin transcript	Axel et al. (1973), 37°C Astrin (1975) + Hg-UTP 37°C	2%
5. Reconstituted chromatin transcript	Astrin (1975) + Hg-UTP 37°C	2%
6. Unsheared chromatin transcript	Astrin (1975) + Hg-UTP 37°C	2%
7. Chromatin transcript	Astrin (1975) + Hg-UTP 25°C	2%
8. Chromatin transcript	Astrin (1975) + Hg-UTP 37°C minus Mn^{++}	2%
9. Endogenous chromatin RNA	--	<0.01%
10. Total L-cell RNA	--	<0.001%

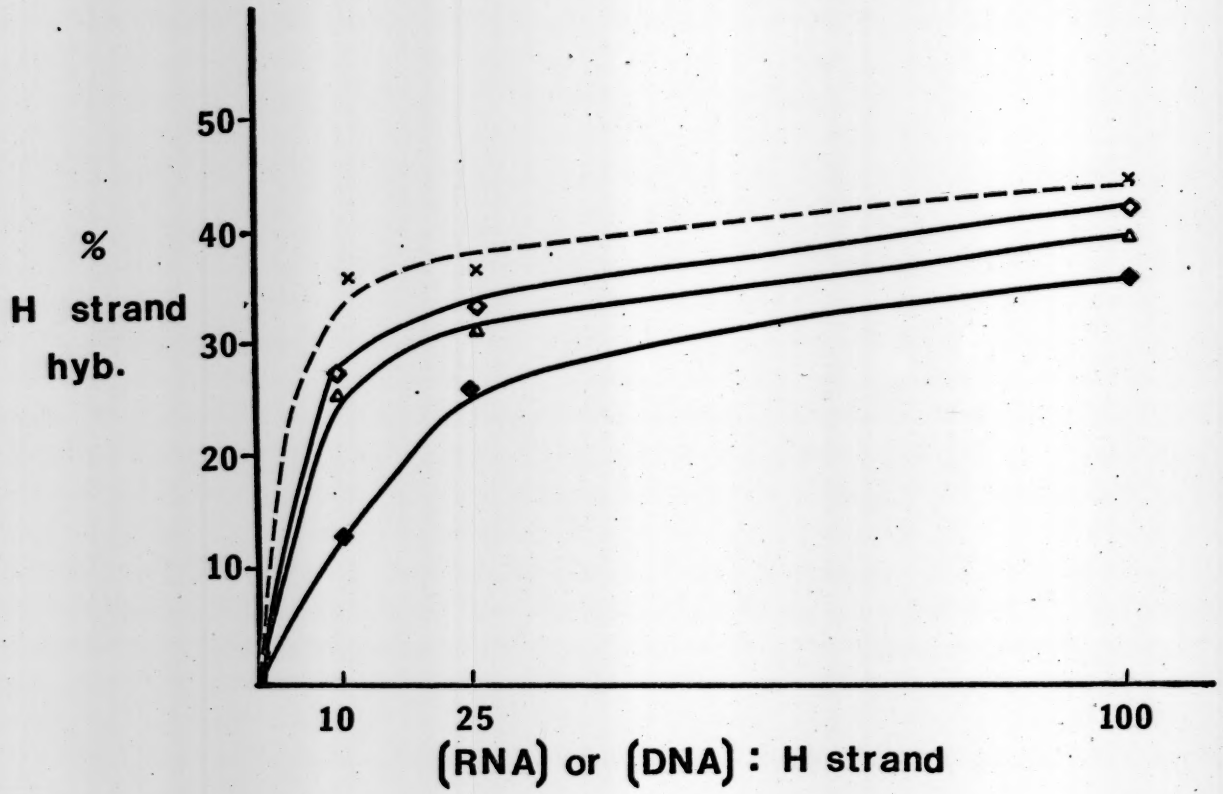
In order to investigate the possibility that the levels of aberrant transcription of satellite sequences were due to the high levels of enzyme used during transcription, mouse chromatin was transcribed using lower enzyme/template ratios (mg/mg). Transcripts were prepared using ratios of 1 to 1, 0.5 to 1, 0.1 to 1, and then were hybridized to satellite DNA. Hybridization of total mouse DNA to satellite DNA was used as a control. Figure 17 shows an experiment in which different amounts of transcript are hybridized to equilibrium with a constant amount of satellite DNA. The initial slopes of the curves give an indication of the proportion of satellite sequences in the RNA or DNA. It is evident from the figure that satellite sequences represent a lower proportion of the sequences in the transcripts than in total DNA and that as the level of enzyme to template is reduced, so also is the proportion of satellite sequences in the in vitro transcript. Therefore it appears that at least some of the aberrant transcription seen in these studies is due to the high levels of enzyme used, a condition which would result in the utilization by the enzyme of weaker, secondary initiation sites such as could be found in satellite regions.

F. Hybridization of Drosophila Chromatin Transcripts to cDNA

In order to obtain a general picture of the fidelity of chromatin transcription by E. coli RNA polymerase, I characterized the in vitro transcription product from Drosophila chromatin or DNA using a cDNA copy to poly(A)-containing nuclear RNA. This probe represents most of the RNA sequences transcribed in Schneider's cells (Levy W. et al., 1976) and therefore offers a general assay for fidelity.

Several hybridization experiments were performed with RNA transcribed in vitro from Schneider's cell chromatin and DNA. Figure 18 shows

Figure 17. Hybridization of ^{3}H-thymidine labeled satellite heavy strand with mouse L-cell DNA or with transcripts synthesized at various ratios of enzyme to template (mg to mg). Various amounts of each transcript preparation or DNA were incubated with a constant amount of heavy strand until the hybridization had reached completion. The amount of H strand in hybrids was then determined by S1 digestion as described in Materials and Methods. Hybridization to DNA (X); to transcript synthesized at enzyme to template ratio of 1 to 1 (◇); transcript synthesized at enzyme to template ratio of 0.5 to 1 (Δ); transcript synthesized at enzyme to template ratio of 0.1 to 1 (◆).

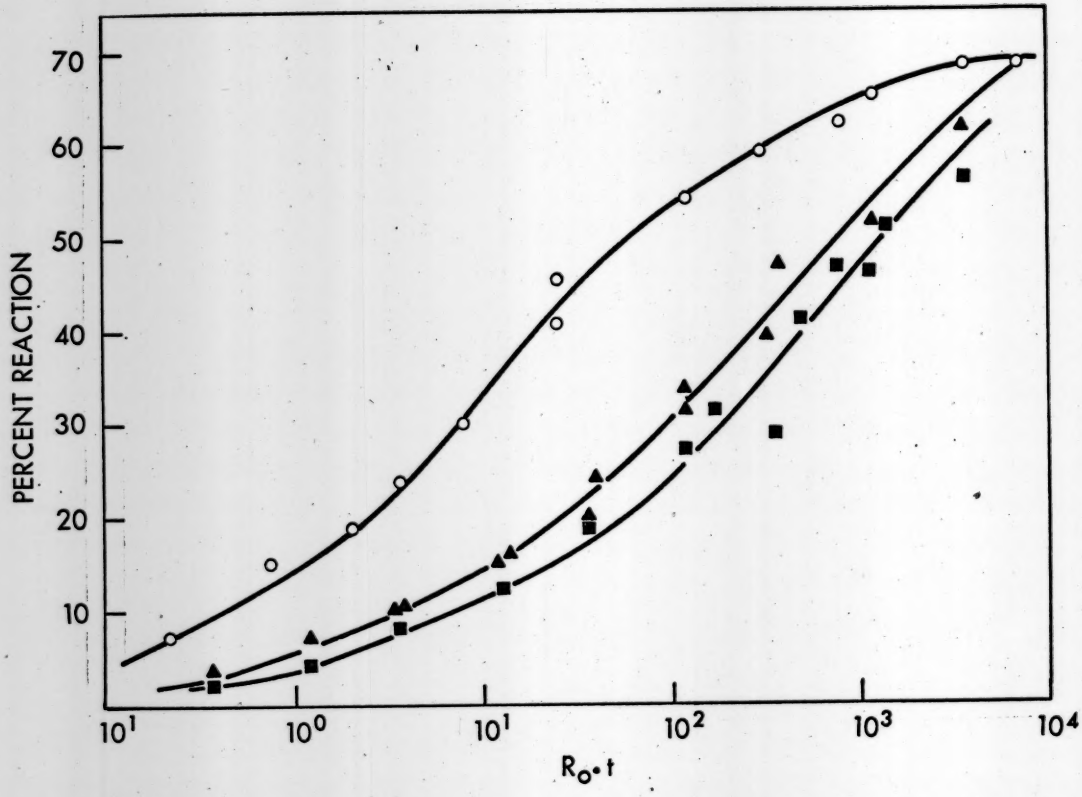


the kinetics of hybridization with the cDNA probe. The kinetics of hybridization of this cDNA probe with its own template is included for comparison. The in vitro transcript from chromatin reacted with the cDNA probe to the same extent as did polyadenylated nuclear RNA, implying that E. coli RNA polymerase had transcribed all of the sequences represented by the probe. However, the reaction of the in vitro transcript was slower by about a factor of 10 compared to the reaction with the in vivo poly(A)-containing nuclear RNA, indicating a dilution of those sequences with others not present in the probe. The relatively small size of the transcript would also tend to produce slower kinetics relative to the in vivo RNA. Taking 6S to be the size of the transcript (see Fig. 14) and 11S to be the size of the poly(A)-containing nuclear RNA (Levy W. et al., 1975), we can calculate that the difference in molecular weight to be about a factor of 10. As shown by Wetmur and Davidson (1971), the rate of hybridization is proportional to molecular weight. Therefore the difference in rates between in vivo and in vitro RNAs could differ by as much as a factor of 3.

The rate of hybridization of the in vitro DNA transcript with the cDNA was displaced by a further factor of two to three. While all sequences present in the probe were transcribed from naked DNA, they make up a smaller proportion of the total than in the case of the chromatin transcript. Thus DNA is a less restricted template than is chromatin.

The complex hybridization kinetics for both in vivo and in vitro transcripts, spanning several log units, implies a broad frequency distribution of sequences in the population. Some sequences both in vivo and in vitro are present much more frequently than are others. The E. coli RNA polymerase seems to select some sequences for transcription

Figure 18. Kinetics of the hybridization of cDNA made from Drosophila polyadenylated nuclear RNA to its own template (O), to RNA synthesized *in vitro* from chromatin (▲) and from DNA (■).


PERCENT REACTION
70
60
50
40
30
20
10
10^1
10^0
10^1
10^2
10^3
10^4
$R_o \cdot t$

more often than others, and it is interesting to ask whether this selectivity is related to *in vivo* abundance. In order to determine whether sequences which are abundant *in vivo* are the same ones which are abundant *in vitro*, a cDNA probe which had been fractionated to represent frequent and infrequent sequences in nuclear RNA was hybridized to *in vitro* transcript. Figure 19 shows the hybridization of these probes with total nuclear RNA. The probe enriched for frequent sequences (frequent cDNA) reacted much more rapidly with nuclear RNA than did the probe which was enriched for infrequent seuqences (infrequent cDNA). This result therefore validates the successful fractionation. Figure 19 also shows the hybridization kinetics obtained when the *in vitro* chromatin transcript was used to drive the two probes. A similar displacement between the two curves was obtained, indicating that *E. coli* polymerase has preferentially transcribed those sequences in chromatin which are also represented by the most abundant RNA in the cell nucleus.

This degree of preferential transcription is not observed when naked DNA is transcribed as illustrated by Figures 20 and 21. Figure 20 shows the hybridization kinetics of total cDNA or frequent cDNA with *in vitro* chromatin transcript. A 10-fold displacement between the two curves is obtained, indicating that the sequences represented by the frequent probe hybridize 10 times faster than the bulk of the sequences in the *in vitro* chromatin transcript. Figure 21 shows the hybridization kinetics of total cDNA or frequent cDNA with *in vitro* DNA transcript. In this case, only a 3-fold displacement between the two curves is obtained, indicating that the sequences represented by the frequent probe hybridize only three times faster than the bulk of the sequences in the *in vitro* DNA transcript, that is, that their frequency is only

Figure 19. Hybridization kinetics of the fractionated cDNA probe to different RNA populations. Probe enriched in frequent sequences hybridized to total Drosophila nuclear RNA (O), to in vitro chromatin transcript (▲); probe enriched in infrequent sequences hybridized to total nuclear RNA (Δ) and to in vitro chromatin transcript (●).

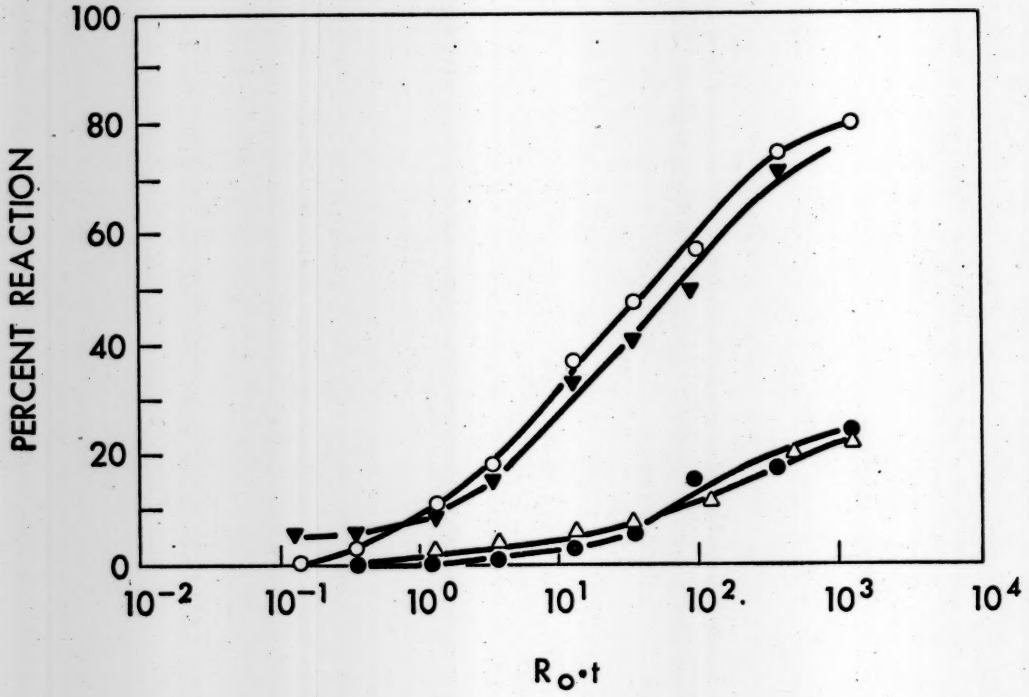

100
80
60
40
20
0
PERCENT REACTION
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
$R_o \cdot t$

Figure 20. Kinetics of hybridization of *in vitro* Drosophila chromatin transcript to cDNA complementary to poly(A)-containing nuclear RNA (●) and to cDNA enriched in frequent nuclear sequences (○).

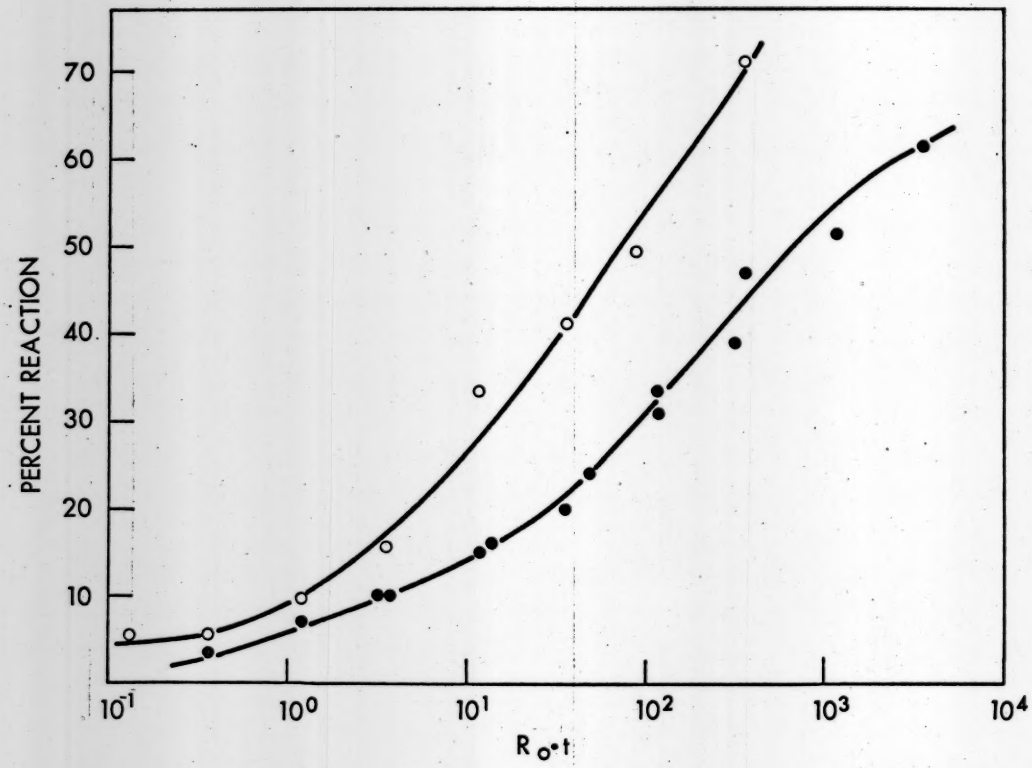
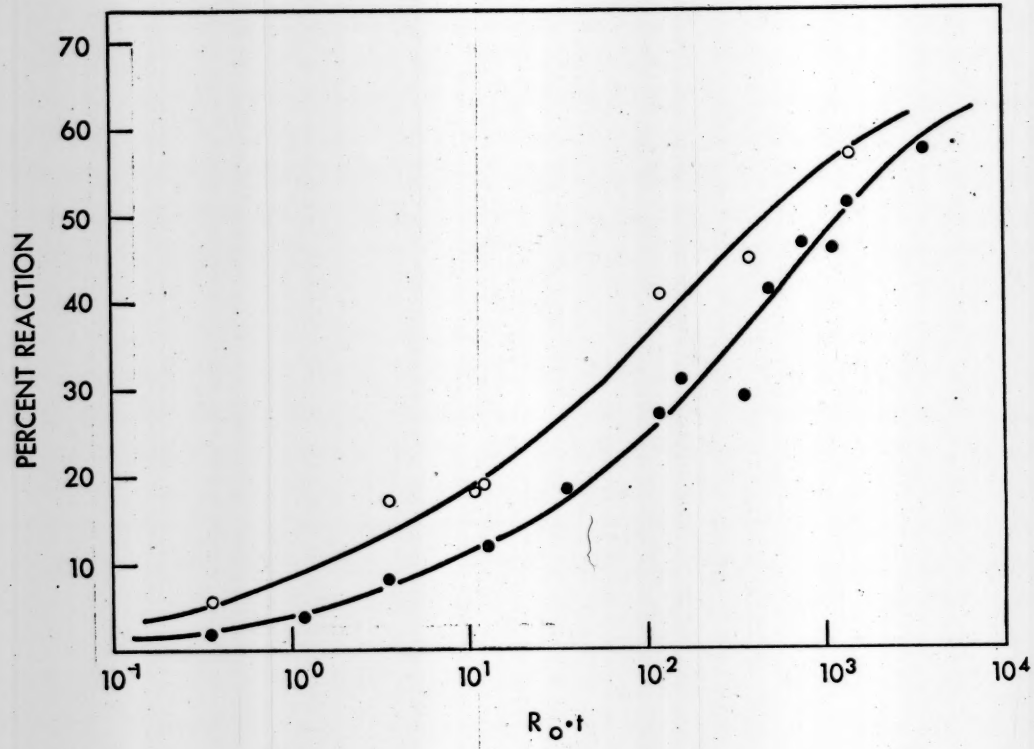

PERCENT REACTION
70
60
50
40
30
20
10
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
$R_o \cdot t$

Figure 21. Kinetics of hybridization of *in vitro* Drosophila DNA transcript to cDNA complementary to poly(A)-containing nuclear RNA (●) and to cDNA enriched in frequent nuclear sequences (○).

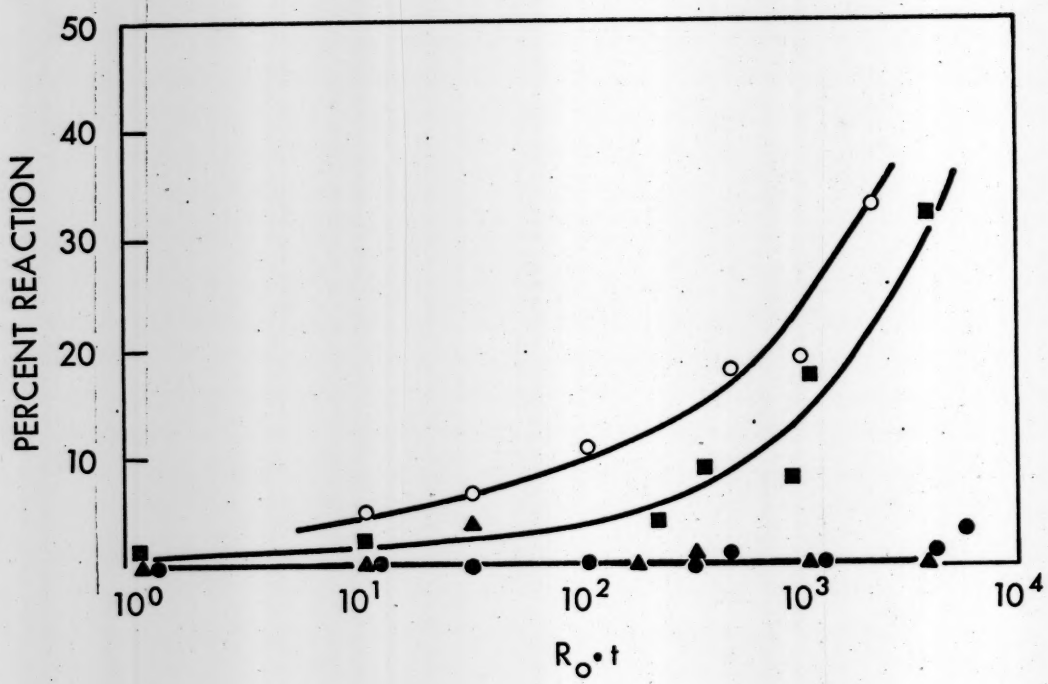

PERCENT REACTION
70
60
50
40
30
20
10
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
$R_o \cdot t$

some 3-fold greater than the average frequency of sequences in the total population. It appears from this result that some vestiges of preferential transcription by E. coli RNA polymerase remain when naked DNA is transcribed, but this can be attributed to gene dosage effects. That is, most of the sequences in the DNA transcript which hybridize rapidly come from sequences which are repeated in the genome. That this is the correct interpretation is shown by comparing the hybridization kinetics of the probes when they are hybridized to DNA (Levy W. and McCarthy, 1976; Levy W. et al., 1976). About a 3-fold displacement of the curves is obtained. The much larger displacement seen in the case of the chromatin transcript cannot be explained in this way, indicating that the chromatin transcript is much more representative of the in vivo frequency distribution than is the DNA transcript.

G. Hybridization of Drosophila Transcripts to Nuclear RNA

In order to examine the symmetry of in vitro transcription, an excess of in vitro transcript was hybridized to labeled total nuclear RNA. As shown in Figure 22, the transcripts from chromatin and DNA both contain sequences which hybridize, that is, sequences from the strand not transcribed in vivo. The observed hybridization of the transcripts to nuclear RNA is not due to self-annealing of the nuclear RNA or to DNA contamination. Due to the low frequency of these wrong strand transcripts in the total RNA population, these hybridizations could not be driven to completion. However, when compared to the hybridization kinetics with the cDNA probe, these curves are shifted to higher $R_o t$ values, suggesting that E. coli polymerase has preferentially transcribed the correct strand. It is possible to make an approximate comparison between the hybridization kinetics for DNA-RNA reactions and

Figure 22. Hybridization kinetics of labeled total Drosophila nuclear RNA to in vitro chromatin transcript (O), in vitro DNA transcript (●), total unlabeled nuclear (▲) and cytoplasmic RNA (●).



RNA-RNA reactions. Bishop et al. (1974) have found that the $R_o t_{1/2}$ for hybridizing α+β globin message (complexity 4×10^5) to its cDNA is 6×10^{-4}. Obinata et al. (1975) have obtained a relationship between the complexity of an RNA and the $R_o t_{1/2}$ for hybridization to its cDNA copy made by Qβ replicase. An RNA species of m.w. 5×10^5 would anneal to its cRNA with a $R_o t$ of about 3×10^{-3}, which is slower than the corresponding DNA-DNA reaction by a factor of 5. Therefore a 20-fold displacement between the cDNA curves and the nuclear RNA curves, actually represents a 4-fold difference in sequence concentrations. Sequences complementary to nuclear RNA are therefore present at about 25% the level of sequences complementary to cDNA. Also included in Figure 22 are control experiments in which the driver was nuclear or cytoplasmic RNA. Very little hybridization occurred, confirming that in vivo RNA is asymmetrically transcribed.

H. Summary of Results

I have shown that all of the sequences transcribed in vivo are also represented in the in vitro transcripts from both chromatin and DNA of Schneider's cells, and furthermore, that the relative abundance of individual sequences is preserved in vitro when chromatin is transcribed but not when DNA is transcribed. It is clear that E. coli RNA polymerase is exerting some preference as to which sequences are transcribed. Nevertheless, some aberrant transcription is detectable. A considerable amount of incorrect strand transcription is evident when Drosophila chromatin is transcribed and some normally repressed satellite sequences are detectable when mouse L-cell chromatin is transcribed.

D. DISCUSSION

The mechanism by which chromosomal components combine with DNA to effect transcriptional regulation is basic to our understanding of gene expression in eucaryotes. Structural and functional studies on chromatin depend on our ability to isolate chromatin in a state which retains the transcriptional regulatory properties of the *in vivo* state. With this in mind, numerous studies have been conducted which attempt to characterize the fidelity of chromatin transcription *in vitro*. The results of these studies leave little doubt that both *E. coli* and mammalian RNA polymerases exert some preference as to which sequences are being transcribed. This conclusion was reached following the demonstration of tissue specific restriction of transcription in isolated chromatins of various organisms using filter hybridization techniques (Paul and Gilmour, 1968; Tan and Miyagi, 1970; Smith *et al.*, 1969; Bekhor *et al.*, 1969; Huang and Huang, 1969). These results were corroborated by subsequent experiments in which the presence or absence of a specific well defined sequence in the *in vitro* transcript was shown to correlate well with the source of the chromatin (Paul *et al.*, 1973; Axel *et al.*, 1973; Steggles *et al.*, 1974; Shih *et al.*, 1975; Astrin, 1973; Astrin, 1975; Stein *et al.*, 1975; Janowski *et al.*, 1974; Smith and Huang, 1976; Harris *et al.*, 1976; Tsai *et al.*, 1976). Thus, for example, globin mRNA sequences are present in the transcript of erythroid but not non-erythroid cell chromatin, and histone sequences are transcribed *in vitro* with an efficiency which depends upon the stage of the cell cycle from which the chromatin is isolated.

Nevertheless, these and the earlier studies suffer from certain limitations in interpretation. The filter hybridization studies were performed under conditions which permit detection only of RNA sequences synthesized from reiterated regions of DNA and present in large amounts, and furthermore they provide only approximate information as to the identity of the transcript. Studies which probe for an identifiable sequence in the transcript, give precise information about the sequence in question, but clearly leave the majority of the transcript uncharacterized. Indeed, it appears that at least some of the uncharacterized portions may contain sequences resulting from aberrant transcription. A considerable amount of wrong strand transcription has been reported (Reeder, 1973; Astrin, 1973; Wilson et al., 1975a,b), some normally non-transcribed spacer is detectable from 5S or ribosomal genes (Reeder, 1973) and some non-transcribed satellite sequences are detectable (Reeder, 1973).

The present experiments were undertaken in an attempt to resolve some of the questions remaining concerning the fidelity of in vitro transcription. A probe which represents a substantial fraction of the sequences produced in vivo (Levy W. et al., 1976), namely, a cDNA copy of Schneider's cell nuclear poly(A)-containing RNA, was used to monitor overall transcriptional fidelity. The infidelity of in vitro transcription was monitored using a nuclear RNA for anti-strand transcription from Drosophila chromatin, and a satellite DNA probe for satellite sequence transcription from L-cell chromatin.

The initial transcription experiments were performed using L-cell chromatin which had been isolated following the method of Bonner et al. (1968) and transcribed following the method of Axel et al. (1973). An

attempt was made to use extremely gentle conditions in the isolation of chromatin and to follow the protocol of Sollner-Webb and Felsenfeld (1975) in which a chromatin gel is formed following an abrupt decrease in ionic strength of the nuclei. These authors report that chromatin prepared in this manner more closely resembles nuclei with respect to its staphylococcal digestion pattern, than does chromatin prepared from nuclei which are slowly swelled. The chromatin was then used within 24 hr for transcription experiments. The transcription conditions which were used kept the salt concentration very low (10 mM Tris, 1 mM $MnCl_2$) in an attempt to minimize protein exchange or sliding. When L-cell chromatin was transcribed under these conditions, a low but reproducible level of satellite transcription occurred, amounting to about one-fourth or one-fifth of the level observed when naked DNA was transcribed (Fig. 5). Since the hybridization kinetics suggest that the DNA transcript is random with respect to satellite (Fig. 4), we can say that satellite sequences make up 10% of the transcript from naked DNA and about 2% of the sequences from chromatin. Consistent with earlier observations (Flamm et al., 1969) no satellite sequences were detectable in cellular RNA (Fig. 5). Less than 0.01% was detectable in RNA endogenous to chromatin and less than 0.001% was detectable in total L-cell RNA.

A technical problem complicating in vitro transcription of chromatin is the presence of DNA and endogenous RNA which co-extracts with the transcript. Since endogenous RNA does not hybridize with satellite DNA, it is possible to deal with this problem by determining the endogenous RNA content in each chromatin preparation and then make the necessary corrections in the hybridization kinetics. However, the presence of DNA in the in vitro transcript is more serious, since it

does contain 10% satellite sequences and will hybridize. Since RNA is preferentially degraded by alkali, this possibility was controlled by digesting a portion of the purified transcript preparation and assaying for remaining hybridizable material. As shown in Figure 5, alkali treatment completely abolished hybridization, indicating that purification of the transcript from DNA had been successful.

It was next of interest to ask whether chromatin could be dissociated and reconstituted and retain the properties of native chromatin following this procedure. Clearly the feasibility of using isolated chromatin for *in vitro* studies of gene regulation depends on our ability to manipulate the composition of the system, and to reconstitute it from its components. Several lines of evidence suggest that reconstitution of chromatin is possible (Bekhor *et al.*, 1969; Huang and Huang, 1969; Paul *et al.*, 1973; Stein *et al.*, 1976; Park *et al.*, 1976) but none of these has dealt with probes for aberrant transcription. In the case of satellite chromatin, one might imagine that higher order coiling resulting from long range interactions would be involved in the mechanism of heterochromatization and transcriptional repression, and that the complex interactions involved in formation of this structure might be difficult to reconstitute. Nevertheless, the results of Figure 5 demonstrate that this is not the case, that satellite sequences reconstitute to the same level of transcriptional repression as existed in the native chromatin.

It was next of interest to generalize the characterization of *in vitro* transcription by asking to what extent the overall population of sequences produced *in vitro* resembled the *in vivo* population. A cDNA copy of nuclear poly(A)-containing RNA from Drosophila cells was used for this purpose. Earlier studies from our laboratory had extensively

characterized the in vivo RNA populations from these cells and had shown that both nuclear and cytoplasmic populations were composed of a complex distribution of species, and that some sequences in the populations were represented much more frequently than were others (Levy W. and McCarthy, 1975, 1976). This finding suggested some interesting questions concerning the frequency distribution of sequences transcribed in vitro. That is, are sequences which are frequent in vivo transcribed more frequently in vitro?

However, in this study certain technical problems had to be considered which had not existed for the earlier studies on L-cell chromatin. First, it was necessary to synthesize in vitro very large amounts of RNA since the hybridization reactions needed to be carried to higher $R_o t$ values. Second, it was necessary to free the in vitro transcript of both contaminating DNA and endogenous RNA. Either contaminant could hybridize to the cDNA and complicate the interpretation of the kinetics. The first problem was resolved by using the transcription system described by Astrin (1975) in which transcripton continues for extended periods of time (up to 4 hr) and produces about 400 µg transcript per mg of chromatin template (Figs. 10 & 11). The salt conditions in this transcription system are higher than in the system of Axel et al. (1975) described above, but this did not seem to affect the level of satellite transcription when L-cell chromatin was transcribed (Fig. 15). There still seemed to be a low level (about 2% of the sequences) of satellite sequences in the transcript.

The chromatin preparation was modified for these experiments. The nuclei were prepared in the presence of sucrose and divalent cations, conditions which tend to stabilize them. They were then subjected to an

abrupt decrease in ionic strength and sheared at 1000 psi. This shearing force is even lower than that used for the chromatin prepared by the method of Bonner et al. (1968) described above.

The second problem which was encountered, the need to remove DNA and endogenous RNA from the in vitro transcript, was resolved by using mercurated UTP (Hg-UTP) in the transcription reaction mixture. As shown by Dale and Ward (1975), this precursor is incorporated into RNA in the presence of a 20-fold excess of 2-mercaptoethanol. The mercaptan is necessary as a mercury ligand. Without it, RNA polymerase will be inactivated. The kinetics of the RNA transcription were only slightly slower as compared to UTP (Fig. 10) and high yields of RNA were still possible. Because of the high specificity of the sulfhydryl column (Table I), this RNA could be completely freed of contaminating endogenous RNA. In resolving this problem, however, one might argue that a new one has been created, that is, the nature of the in vitro population containing the mercurated precursor may not be the same as the population using the unmercurated precursor. It may be that the Km for the mercurated substrate is lower, causing the enzyme to select against sequences which contain large amounts of uridine in favor of sequences enriched in the other bases. This is probably not the case. Beebee and Butterworth (1976) have shown that the Km's for Hg-UTP and UTP are comparable for RNA polymerase I from rat liver and moreover that the base compositions of in vitro transcripts are the same for in vitro transcripts synthesized in the presence of UTP or Hg-UTP.

Therefore, several experiments on mouse L-cell chromatin were conducted to compare this set of procedures for chromatin isolation and transcription, with the earlier set. Transcription was performed at

25°C (as was Drosophila chromatin transcription) or at 37°C. Regulatory proteins on chromatin might be expected to show a temperature dependence. As shown in Figure 15, the level of satellite sequence expression was the same in both cases. As shown in Figure 16, when Mn^{++} was omitted from the reaction mixture, no decrease in satellite sequence expression was observed. Therefore it cannot be argued that Mn^{++} induces aberrant initiations. Also shown in Figure 16 is an experiment in which the transcript from unsheared chromatin is hybridized to satellite DNA. In this case, nuclei are lysed in 10 mM Tris and transcribed immediately. Levels of transcription are much lower in this case, but are sufficient to provide enough transcript for the hybridization analysis. Again, a low but significant level of satellite transcription is observed, equivalent to the amount observed from sheared chromatin. However, as shown in Figure 17, when L-cell chromatin was transcribed using lower ratios of enzyme to template, the proportion of satellite sequences in the transcript was reduced. This suggests that the *in vitro* transcription from chromatin regions which are transcriptionally inactive *in vivo*, is at least in part an artifact caused by the use of saturating levels of enzyme. Under limiting enzyme conditions, satellite regions compete poorly with other regions of the chromatin for available enzyme. Since satellite regions are selected against under these conditions, they must offer only weak initiation sites for transcription.

Experiments were then conducted using Drosophila chromatin as template and cDNA complementary to poly(A)-containing nuclear RNA as probe. In doing this, a 10-fold shift of the $R_o t_{1/2}$ values was observed between the hybridization curve of this probe to *in vivo* poly(A)-containing nuclear RNA and RNA synthesized *in vitro* from a chromatin template.

Although most of the sequences present in the poly(A) nuclear RNA population were also transcribed *in vitro*, they appear to be diluted with other sequences not represented in the probe. However, the use of the hybridization curve of the cDNA to its own template as reference presents several difficulties. First, it does not take account of nuclear sequences lacking poly(A) which could also be synthesized *in vitro*. Second, it must be stressed that the abundance of different RNA molecules in the nucleus reflects a balance between the rates of biosynthesis and of processing. Since processing enzymes are probably not isolated with chromatin, it is likely that the relative abundance of different sequences at the end of the *in vitro* transcription depends only upon the relative efficiency with which they are produced by the *E. coli* polymerase. Therefore, the hybridization kinetics of *in vitro* and *in vivo* RNAs with the same cDNA probe need not be the same. It can only be concluded that all the sequences represented in the nuclear cDNA probe can be synthesized *in vitro* and that the proportion of the total population which they represent is about three times higher for the chromatin transcript than for that from DNA. If we assume that the transcript from DNA is essentially random, then it seems to follow that in this case transcription from chromatin as template is only some three times better than random.

More meaningful questions, however, can be asked regarding the frequency distribution of sequences produced *in vitro*. In the case of *in vivo* populations, it is clear that some sequences in both nucleus and cytoplasm are present much more frequently than are others (see above). Therefore it is important to know whether chromosomal sites which code for the most abundant RNA in the cell nucleus are preferentially transcribed in the chromatin system. The use of a cDNA probe fractionated to

represent the most abundant sequences in nuclear RNA allows one to approach this question. This experiment yielded a very positive result in that the rate of hybridization with abundant cDNA was some hundred fold greater than for the probe representing rarer sequences. Therefore it is clear that different sites in chromatin are transcribed with relative efficiencies which differ by two orders of magnitude, and it appears that this difference in efficiency is related to in vivo abundance. However, the formal possibility cannot be eliminated that relative abundance is controlled predominantly by processing rather than synthesis in vivo and that this correlation between in vivo and in vitro abundance is a fortuitous result of aberrant transcription. Nevertheless, this result, together with the results concerning the transcription of specific RNA sequences from chromatin (Paul et al., 1973; Axel et al., 1973; Shih et al., 1973; Astrin, 1973; Reeder, 1973; Stein et al., 1975) implies that the E. coli enzyme can recognize some features of chromatin structure which directly relate to preferential transcription in the cell. That this is attributable to chromatin structure rather than DNA base sequence is demonstrated by the failure of the DNA transcript to drive abundant cDNA to a degree greater than that explicable by gene dosage effects.

A certain low level of aberrant transcription is observed when either L-cell chromatin or Drosophila chromatin is transcribed with E. coli RNA polymerase. In the case of isolated L-cell chromatin, satellite sequence expression from chromatin is about 20-25% of the level from naked DNA. When Drosophila chromatin is transcribed, a rough calculation suggests that the wrong strand is expressed at about 25% of the level of the correct strand (see Results). This result is in agreement with other reports of poor strand selection or transcription of normally

non-transcribed regions (Reeder, 1973; Honjo and Reeder, 1974; Astrin, 1973; Wilson et al., 1975b).

How can this set of conflicting reports be reconciled. On the one hand, using cDNA complementary either to individual mRNA's or to the overall population of polyadenylated RNA, it does appear that selective transcription occurs at some sites which are related to activities in the cell or tissue from which the chromatin originated. On the other hand, there is no reason to believe that the E. coli enzyme correctly recognizes initiation and termination signals. Indeed, binding of E. coli enzyme and homologous enzyme are non-competitive (Cedar, 1975). In addition, there can be little doubt that incorrect strand selection and other types of aberrant transcription occur. The most straightforward explanation is that the non-homologous polymerase can recognize some aspect of chromatin structure which is preserved during isolation and correlated with transcriptional activity. The fact that the E. coli enzyme can, in addition, preferentially transcribe sequences which are abundant in vivo, suggests that the transcriptional capacity of a given region in chromatin is determined by its structure; that is, the physical availability of a sequence is directly related to the frequency with which it is recognized by foreign or homologous enzymes. Thus, the more frequently transcribed sequences would be in a more open conformation. This probably explains why the E. coli enzyme, despite its probable failure to recognize eucaryotic control elements or initiation signals, has been such a useful probe for transcriptional specificity. The abilities of other DNA binding proteins, DNase I and DNase II, to attack active chromatin preferentially, would be analogous findings (Weintraub and Groudine, 1976; Gottesfeld et al., 1974).

In conclusion, these results show that the components of chromatin serve to restrict the transcription of DNA in a very specific manner, and that isolated chromatin can be transcribed with a high degree of fidelity *in vitro*. In addition, it appears that *E. coli* RNA polymerase is able to recognize the correct sequences for transcription on both a qualitative and quantitative level. Some of the aberrant transcription which is observed seems to be due to binding of enzyme to weak initiation sites, and can be minimized by using lower levels of enzyme. However, it may be that another part of the aberrant transcription results from the failure of the heterologous enzyme to recognize correct initiations and terminations on eucaryotic templates and from possible disruption of critical interactions between chromosomal components during chromatin isolation. Nevertheless, it appears that most of the essential regulatory elements of chromatin are recognized *in vitro*, a finding which supports the feasibility of further studies on chromatin structure and transcriptional regulation.

E. REFERENCES

Astrin, S. (1973). Proc. Nat. Acad. Sci. USA 70, 2304.

Astrin, S. (1975). Biochemistry 14, 2700.

Axel, R., Cedar, H., Felsenfeld, G. (1973). Proc. Nat. Acad. Sci. USA 70, 2029.

Axel, R., Feigelson, P. & Schultz, G. (1976). Cell 7, 247.

Beebee, T.J.C. & Butterworth, P.H.W. (1976). Eur. J. Biochem. 66, 543.

Bekhor, I., Kung, G.M. & Bonner, J. (1969). J. Mol. Biol. 39, 351.

Bishop, J.O., Morton, J.G., Rosbash, M. & Richardson, M. (1974). Nature 250, 199.

Bonner, J., Chalkley, G.R., Dahmus, M., Fambrough, D., Fugimura, F., Huang, R.C., Huberman, J., Jensen, R., Marushige, K., Ohlenbusch, H., Olivera, B. & Widholm, J. (1968). Methods in Enzym. 12b, 3.

Bonner, J., Dahmus, M.E., Fambrough, D., Huang, R.C.C., Marushige, K., & Tuan, D.Y.H. (1968). Science 159, 47.

Brown, I.R. & Church, R.B. (1972). Develop. Biol. 29, 73.

Burgess, R.R. (1969). J. Biol. Chem. 244, 6160.

Campo, M.S. & Bishop, J.O. (1974). J. Mol. Biol. 90, 649.

Cedar, H. (1975). J. Mol. Biol. 95, 257.

Cedar, H. & Felsenfeld, G. (1973). J. Mol. Biol. 77, 237.

Chae, C.-B. (1975). Biochemistry 14, 900.

Cox, R.F. (1976). Cell 7, 455.

Cuatrecasas, P. (1970). J. Biol. Chem. 245, 3059.

Dale, R.K.M. & Ward, D.C. (1975). Biochemistry 14, 2458.

Davidson, E.H. & Hough, B.R. (1971). J. Mol. Biol. 56, 491.

Echalier, G. & Ohanessian, A. (1970). In Vitro 6, 162.

Firtel, R.A. (1972). J. Mol. Biol. 66, 363.

Flamm, W.G., Walker, P.M.B. & McCallum, M. (1969). J. Mol. Biol. 40, 423.

Gelderman, A.H., Rake, A.V. & Britten, R.J. (1971). Proc. Nat. Acad. Sci. USA 68, 172.

Getz, M.J., Birnie, G.D., Young, B.D., MacPhail, E. & Paul, J. (1975). Cell 4, 121.

Gilmour, S. & Paul, J. (1973). Proc. Nat. Acad. Sci. USA 70, 3440.

Gottesfeld, J.M., Garrard, W.T., Bagi, G., Wilson, R.F. & Bonner, J. (1974). Proc. Nat. Acad. Sci. USA 71, 2193.

Grady, L.J. & Campbell, W.P. (1973). Nature New Biol. 243, 195.

Grouse, L., Chilton, M.D. & McCarthy, B.J. (1972). Biochemistry 11, 798.

Gurdon, J.B. (1962). Develop. Biol. 4, 256.

Hahn, E.W. & Laird, C.D. (1971). Science 173, 158.

Hager, G., Holland, M., Valenzuela, P., Weinberg, F. & Rutter, W.J. (1976). in "RNA Polymerase" (Losick, R. & Chamberlin, M., eds) Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. p. 745.

Harris, S.E., Swartz, R.J., Tsai, M.J., O'Malley, B.W. & Roy, A.K. (1976). J. Biol. Chem. 251, 524.

Hewish, D.R. & Burgoyne, L.A. (1973). Biochem. Biophys. Res. Commun. 52, 504.

Honjo, T. & Reeder, R.H. (1974). Biochemistry 13, 1896.

Huang, R.C.C. & Huang, P.C. (1969). J. Mol. Biol. 39, 365.

Hunt, J.A. (1974). Biochem. J. 138, 499.

Janowski, M., Baugret-Mahieu, L. & Sassen, A. (1974). Nature 251, 347.

Kakpakov, V.T., Gvosdev, V.A., Platova, T.P. & Polukarova, L.C. (1969). Genetika 5, 67.

Kennell, D. (1968). J. Mol. Biol. 34, 85.

Kleiman, L. & Huang, R.C.C. (1972). J. Mol. Biol. 64, 1.

Laird, C.D. (1971). Chromosoma 32, 378.

Laird, C.D., Chooi, W.Y., Cohen, E.H., Dickson, E., Hutchinson, N. & Turner, S.H. (1973). Cold Spring Harbor Symp. Quant. Biol. 38, 311.

Leong, J.A., Garapin, A.C., Jackson, N., Fanisher, L., Levinson, W. & Bishop, J.M. (1972). J. Virol. 9, 891.

Levy, W., B., Johnson, C.B., McCarthy, B.J. (1976). Nucl. Acids Res. 3, 177.

Levy, W., B., & McCarthy, B.J. (1975). Biochemistry 14, 2440.

Levy, W., B., & McCarthy, B.J. (1976). Biochemistry 15, 2415.

Lieberman, M.W. (1973). Biochim. Biophys. Acta 324, 309.

Lindsley, D.L. & Grell, E.H. Genetic Variations of Drosophila Melanogaster, Carnegie Inst. of Washington Publ. No. 627 (1968, 1972).

Maio, J.J. & Kurmit, D.M. (1974). Biochim. Biophys. Acta 349, 305.

Marzluff, W.F. & Huang, R.C.C. (1975). Proc. Nat. Acad. Sci. USA 72, 1082.

McCarthy, B.J. & Bolton, E.T. (1964). J. Mol. Biol. 8, 184.

McCarthy, B.J. & Hoyer, B.H. (1964). Proc. Nat. Acad. Sci. USA 52, 915.

McCarthy, B.J., Nishiura, J.T., Doenecki, D., Nasser, D.S. & Johnson, C.B. (1973). Cold Spring Harbor Symp. Quant. Biol. 38, 763.

McConaughy, B., Laird, C.D. & McCarthy, B.J. (1969). Biochemistry 8, 3289.

Melli, M., Whitfield, C., Rao, K.V., Richardson, M. & Bishop, J.O. (1971). Nature New Biol. 231, 8.

Mirsky, A.E. & Ris, H. (1951). J. Gen. Physiol. 34, 451.

Obinata, M., Nasser, D.S. & McCarthy, B.J. (1975). Biochem. Biophys. Res. Commun. 64, 640.

Palmiter, R.D. (1974). Biochemistry 13, 3606.

Pardue, M.L., Gerbi, S.A., Eckhardt, R.A. & Gall, J.G. (1970). Chromosoma 29, 268.

Park, W.D., Stein, J.L. & Stein, G.S. (1976). Biochemistry 15, 3296.

Paul, J., & Gilmour, R.S. (1968). J. Mol. Biol. 34, 305.

Paul, J., Gilmour, R.S., Affara, N., Birnie, G., Harrison, A., Hell, S., Humphries, S., Windass, J. & Young, B. (1973). Cold Spring Harbor Symp. Quant. Biol. 38, 885.

Peacock, W.J., Brutlag, D., Goldring, E., Appels, R., Hinton, C.W. & Lindsley, D.L. (1973). Cold Spring Harbor Symp. Quant. Biol. 38, 405.

Randerath, E. & Randerath, K. (1964). J. Chrom. 16, 126.

Rasch, E.M., Barr, H.J. & Rasch, R.W. (1971). Chromosoma 33, 1.

Reeder, R.H. (1973). J. Mol. Biol. 80, 229.

Rothfels, K., Sexsmith, E., Heimburger, M. & Daruse, M.O. (1966). Chromosoma 20, 54.

Rudkin, G.T. & Tartof, K.D. (1973). Cold Spring Harbor Symp. Quant. Biol. 38, 397.

Ryffel, G.U. & McCarthy, B.J. (1975). Biochemistry 14, 1379.

Schneider, I. (1972). J. Embr. Exp. Morph. 27, 353.

Schneider, I. & Blumenthal, A.B. (1976). In: The Genetics and Biology of Drosophila, vol. 2 (T.R.F. Wright & M. Ashburner, eds). Academic Press, N.Y. (in press).

Shapiro, H.S. (1970). In: Handbook of Biochemistry (H.H. Sober, ed), Chemical Rubber Co., Cleveland, Ohio, p. 104.

Shih, T.Y., Khoury, G. & Martin, M.A. (1973). Proc. Nat. Acad. Sci. USA 70, 3506.

Smith, K.D., Church, R.B. & McCarthy, B.J. (1969). Biochemistry 8, 4271.

Smith, M.M. & Huang, R.C.C. (1976). Proc. Nat. Acad. Sci. USA 73, 775.

Sollner-Webb, B. & Felsenfeld, G. (1975). Biochemistry 14, 2915.

Steggles, A.W., Wilson, G.N., Kantor, J.A., Picciano, D.J., Falvey, A.K. & Anderson, W.F. (1974). Proc. Nat. Acad. Sci. USA 71, 1219.

Stein, G.S., Park, W., Thrall, C.L., Mans, R.J. & Stein, J.L. (1975). Nature 257, 764.

Stein, J.L., Reed, K. & Stein, G.S. (1976). Biochemistry 15, 3291.

Sutton, W. (1971). Biochim. Biophys. Acta 240, 522.

Tan, C.H. & Miyagi, M. (1970). J. Mol. Biol. 50, 641.

Tsai, S.Y., Harris, S.E., Tsai, M.J. & O'Malley, B.W. (1976). J. Biol. Chem. 251, 4713.

Turner, S.H. & Laird, C.D. (1973). Biochem. Genet. 10, 263.

van Keulen, H., Planta, R.J. & Retel, J. (1975). Biochim. Biophys. Acta 395, 179.

Vendreley, R. (1955). The Nucleic Acid, vol. 2, p. 155.

Waring, M. & Britten, R.J. (1966). Science 154, 791.

Weintraub, H. & Groudine, M. (1976). Science 193, 848.

Wetmur, J.G. & Davidson, N. (1968). J. Mol. Biol. 31, 349.

Wilson, G.N., Steggles, A.W., Kantor, J.A., Nienhuis, A.W. & Anderson, W.F. (1975a). J. Biol. Chem. 250, 8604.

Wilson, G.N., Steggles, A.W. & Nienhuis, A.W. (1975b). Proc. Nat. Acad. Sci. USA 72, 4835.

Wu, J.R., Hurn, J. & Bonner, J. (1972). J. Mol. Biol. 64, 211.